LAW / TECHNOLOGY / INSTITUTIONS

Artificial Intelligence and Legal Transformation

IN THE GLOBAL SOUTH

Comparative legal analysis and institutional design



Muhammad Faiq Butt
Muhammad Sohail Asghar
Hafsa Naz

LUMINA LITERATI PUBLISHING 2026

LUMINA LITERATI PUBLISHING

Artificial Intelligence and Legal Transformation in the Global South

Comparative legal analysis and institutional design

Muhammad Faiq Butt

Muhammad Sohail Asghar

Hafsa Naz

2026

PUBLICATION & RIGHTS

Artificial Intelligence and Legal Transformation in the Global South

© 2026 Muhammad Faiq Butt, Muhammad Sohail Asghar and Hafsa Naz.

The authors retain 100% of the copyright in this work.

Published in 2026 by
Lumina Literati Publishing
luminaliterati.com

ISBN: 978-627-7813-67-3
DOI: 10.65758/isbn.9786277813673

OPEN ACCESS · CC BY-NC-ND 4.0

This book is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. Noncommercial sharing of the unchanged work is permitted with appropriate attribution and a link to the licence. Adapted versions may not be distributed under this licence. Uses permitted by law are unaffected.

creativecommons.org/licenses/by-nc-nd/4.0/

SUGGESTED CITATION

Butt, M. F., Asghar, M. S., & Naz, H. (2026). *Artificial Intelligence and Legal Transformation in the Global South*. Lumina Literati Publishing. https://doi.org/10.65758/isbn.9786277813673

Cover artwork generated with OpenAI GPT Images.

Contents

About this book iv
Abbreviations and short titles vi
About the authors vii
List of tables and models x
Chapter 1 · Legal Systems and the Conditions of AI Adoption 1
1.1 AI Systems and the Meaning of Legal Transformation 3
1.2 Legal Pluralism and Institutional Variation 12
1.3 Legal Authority for Adopting and Using AI 19
1.4 Comparative Research Design and Source Assessment 26
1.5 Criteria for Evaluating Legal and Institutional Change 32
Chapter 2 · AI in Legal Practice and the Administration of Justice 43
2.1 Legal Research, Drafting, and Document Review 46
2.2 Court Administration and Case Management 54
2.3 Judicial Reasoning and AI-Assisted Decision-Making 61
2.4 Legal Aid, Language Access, and Public Legal Information 69
2.5 Professional Duties and Changes in Legal Work 77
Chapter 3 · Data Governance and Automated Decisions 88
3.1 Legal Bases for Data Collection and Reuse 90
3.2 Legal Data, Language Coverage, and Intellectual Property 97
3.3 Public Benefits and Identity: Historical Administrative Analogues for AI 104
3.4 Automated Decisions in Employment, Credit, and Digital Platforms 112
3.5 Cross-Border Processing and Technology Supply Arrangements 119
Chapter 4 · Liability, Evidence, and Legal Remedies 129
4.1 Allocation of Responsibility Across AI Systems 131
4.2 Notice, Reasons, and Access to Decision Records 140
4.3 AI-Related Evidence and Standards of Proof 147
4.4 Individual Review and Corrective Remedies 156
4.5 Collective Harm and Institutional Enforcement 163
Chapter 5 · Regulation, Procurement, and Institutional Implementation 171
5.1 Regulatory Instruments and Institutional Mandates 173
5.2 Procurement Requirements and Contractual Safeguards 178
5.3 Impact Assessment and Local Validation 186
5.4 Staffing, Training, and Financial Planning 194
5.5 Evaluation, Incident Response, and System Withdrawal 205
Conclusion 214
References 217

About this book

Artificial intelligence changes legal institutions through the uses to which its outputs are put. This book examines when assistance becomes consequential control, how legal systems distribute that control, and what an affected person needs to contest its exercise. It addresses scholars of law and technology, lawyers, courts, public authorities and those responsible for designing or procuring legal services.

The argument combines selected domestic legal comparisons with a practical account of implementation. The comparisons establish conditions of authority, representation and remedy; the instruments make those conditions usable in an institutional decision. Neither strand stands alone. A technically complete safeguard can fail where no official has the power or time to act on it. A sound statement of law can fail where the record needed to invoke it cannot be recovered.

The book's central conclusion is that accountable legal transformation requires an institution to connect each consequential computational use to lawful responsibility, usable evidence and effective correction. This does not require every process to become digital or every uncertainty to be eliminated. It requires a defensible boundary around the task, a reasoned choice among feasible alternatives and a route through which errors can change the result.

How to read and use the book

Chapter 1 supplies the thesis, comparative method and evaluation framework. Chapter 2 examines professional and judicial work, including a paired comparison of court rules and an integrated protocol exercise. Chapter 3 follows data through collection, reuse, transfer and correction. Chapter 4 examines evidence, liability and remedies through worked disputes. Chapter 5 connects procurement, staffing, security, evaluation and exit. The conclusion draws together what these analyses establish and what remains to be tested in practice.

The institutional forms and model clauses are proposals unless an identified legal source makes the relevant requirement binding. Constructed examples are labelled where introduced; their facts and numerical assumptions illustrate reasoning, not observed deployments. Domestic legal arguments identify their applicable sources and limits. Section 1.4 explains the selection and verification method, including its language and coverage boundaries. References are grouped into cases, legislation and institutional instruments, scholarship, and technical or policy materials, with each underlying work listed once.

One record and five linked inquiries

Create one master record for a proposed application. Give it an application identifier and retain its purpose, institution, legal function, permitted consequence, responsible office, source collection, system version, approval date and review date. Individual matters receive separate identifiers linked to that record. Dataset, processing activity, test and incident records should refer to the same identifiers where appropriate. Protected personal information belongs in access-controlled records, not in a public project register.

Stage and instruments	Question resolved	Record to carry forward
Authority and alternatives: Tables 1.1-1.5	What task and consequence are proposed, under whose power, and compared with what alternative?	Approved scope, legal basis, source boundaries, comparator and unresolved conditions
Professional workflow: Chapter 2 instruments	Who checks sources, applies judgment and approves consequential work?	Version, evidence, reviewer, departures and released output
Data lifecycle: Chapter 3 instruments	What information may be used, disclosed, corrected or retained?	Dataset and activity identifiers, permissions, transfer basis and correction recipients
Dispute and remedy: Chapter 4 instruments and Models 4.1-4.2	Can the affected person reconstruct, challenge and change the result?	Matter chronology, evidence hold, responsible decision-maker, review route and implemented relief
Delivery and continuity: Chapter 5 instruments and clauses 5.2.1-5.2.12	Can the institution sustain the controls, respond to incidents and exit?	Test and acceptance record, staff assignments, change history, costs and surviving obligations

The core record is maintained by a designated custodian; legal approvals remain with the actors empowered to give them. Specialist modules should refer to shared facts rather than repeat them. If two modules disagree, preserve the discrepancy, identify the authoritative source and record the correction with its date and owner. Do not silently overwrite the earlier decision. A small institution can begin with the core scope and authority record, a proportionate data record, a release-and-review log, and an incident-and-correction route. Additional modules follow the task's consequences, not a requirement to complete every table mechanically.

The Chapter 2 protocol exercise demonstrates this shared-record approach. Chapter 4's Case M follows a constructed false identity match through legal responsibility and relief; Chapter 5 returns to its intake and capacity implications. These exercises are complementary: one connects the workflow instruments, while the other follows an error from identification to correction. The assignment worksheet in section 5.4 maps functions to people without assuming that every function requires a separate office.

Abbreviations and short titles

Abbreviation	Meaning
AI	Artificial intelligence
ANPD	Brazilian data protection authority
AU	African Union
BISP	Benazir Income Support Programme
BSA	Bharatiya Sakshya Adhiniyam 2023
CDC	Brazilian Consumer Defence Code
CNJ	Conselho Nacional de Justiça
LGPD	Brazilian General Personal Data Protection Law
NIST	National Institute of Standards and Technology
OECD	Organisation for Economic Co-operation and Development
PAJA	Promotion of Administrative Justice Act 3 of 2000
POPIA	Protection of Personal Information Act 4 of 2013
RMF	Risk Management Framework
UNESCO	United Nations Educational, Scientific and Cultural Organization

About the authors

Muhammad Faiq Butt

faiqbutt55@gmail.com

Muhammad Faiq Butt is an Assistant Professor of Law at the University of Okara, with almost twelve years of teaching and research experience in legal education. His academic specialization is in criminal law, providing a strong foundation for examining emerging technologies, legal responsibility, institutional accountability, and the administration of justice.

His contribution to Artificial Intelligence and Legal Transformation in the Global South reflects his interest in how technological developments are reshaping legal institutions, regulatory practices, and enforcement mechanisms. His expertise in criminal law offers valuable insights into the implications of artificial intelligence for policing, automated decision-making, evidentiary processes, and legal responsibility.

He also attended a 15-day international training program in China titled “Future-Ready Academic Leadership: Governance, Innovation, and AI Integration in Higher Education,” hosted by Tianjin Foreign Studies University, China. Within the book's comparative framework, he contributes to interdisciplinary analysis of technological change across Global South jurisdictions.

About the authors

Muhammad Sohail Asghar

muhammad.s.asghar@gmail.com
ORCID: 0009-0002-5317-4920

Muhammad Sohail Asghar is a Tenured Associate Professor of Law with more than fourteen years of teaching and research experience and an academic background in law spanning three countries: Pakistan, the United Kingdom, and Malaysia. His international legal education provides a comparative foundation for examining diverse legal systems and the evolving relationship between law and technology.

His research focuses on the intersection of law and technology, particularly the legal, constitutional, and institutional challenges arising from artificial intelligence and emerging digital technologies. In Artificial Intelligence and Legal Transformation in the Global South, he contributes a comparative and interdisciplinary perspective to regulatory diffusion, legal adaptation, institutional implementation, and transnational AI governance.

His scholarly interests encompass AI regulation, privacy, equality, legal accountability, and the transformation of legal frameworks in response to technological developments across diverse legal and institutional contexts.

About the authors

Hafsa Naz

hnhafsa777@gmail.com
ORCID: 0009-0008-9426-2429

Hafsa Naz is an academician and technology researcher with an educational background in computer science and almost a decade of teaching and research experience. She has participated in numerous interdisciplinary national and international conferences, contributing to academic dialogue across technology and related fields.

Her research focuses on the development and deployment of artificial intelligence technologies, including their technical specifications, system architecture, and operational characteristics. With her foundation in computer science, she brings a technical and interdisciplinary perspective to the study of AI-related legal transformation.

In Artificial Intelligence and Legal Transformation in the Global South, her expertise is particularly relevant to examining AI infrastructure, technical standards, system deployment, and technological dependencies that shape institutional and legal responses. Her contribution supports an integrated understanding of how technical capabilities, implementation conditions, and regulatory requirements interact across diverse Global South jurisdictions.

List of tables and models

Table 1.1. System functions and possible legal consequences 11
Table 1.2. Institutional authority and knowledge map 18
Table 1.3. Legal-authority assessment form 25
Table 1.4. Comparative case-selection and evidence-extraction template 30
Table 1.5. Evaluation matrix for a proposed legal AI application 39
Table 2.1. Pre-release protocol for AI-assisted legal work 52
Table 2.2. Court-administration workflow for consequential steps 59
Table 2.3. Task-based protocol for judicial AI assistance 66
Table 2.4. Public legal-information and legal-aid service protocol 75
Table 2.5. Professional-use policy and allocation of review responsibilities 83
Table 3.1. Data-use register: activities, permissions, custody, and affected-person procedures 95
Table 3.2. Corpus assessment for legal authority, language, privacy, and permitted use 102
Table 3.3. Decision record connecting eligibility, evidence, output, and official action 110
Table 3.4. Sector comparison of private automated decisions and routes of protection 117
Table 3.5. Cross-border dependency, evidence-access, and continuity register 126
Table 4.1. Responsibility across an AI-assisted decision process 138
Model 4.1. Notice of a decision involving an AI system 144
Model 4.2. Request for decision records and preservation 144
Table 4.2. Preservation requirements for an AI-related dispute 153
Table 4.3. Questions for an expert examining AI-related evidence 155
Table 4.4. From complaint to effective individual relief 162
Table 4.5. Investigation, structural correction, and monitoring 168
Table 5.1. Regulatory instruments, implementing bodies, and enforcement questions 177
Table 5.2. Task-specific local validation record 189
Table 5.3. Impact-assessment and pilot-authorisation template 193
Table 5.4. Role and responsibility schedule for institutional implementation 196
Table 5.5. Lifecycle cost and dependency register 200
Table 5.6. Phased implementation plan 204
Table 5.7. Model review calendar 206
Table 5.8. Institutional incident-response protocol 208
Table 5.9. Continuation, restriction, suspension, or withdrawal decision form 211

CHAPTER

01

Legal Systems and the Conditions of AI Adoption

IN THIS CHAPTER

1.1 AI Systems and the Meaning of Legal Transformation

1.2 Legal Pluralism and Institutional Variation

1.3 Legal Authority for Adopting and Using AI

1.4 Comparative Research Design and Source Assessment

1.5 Criteria for Evaluating Legal and Institutional Change

Introduction

Legal practice can be changed at an institution level without a legal change, for example a law could continue to exist and yet the way in which a legal authority is able to exercise its legal powers can be changed without changing the law itself. This is particularly true in the area of finding legal authorities. A searchable database may make it easier for a lawyer to find an applicable legal precedent. However, a system that controls which information reaches a decision-maker, routes a complaint to a particular procedure, or generates reasons for denying an application may change how existing authority is exercised. Therefore, the key point of difference is not whether an institution uses AI or not. Rather, it is based on the type of legal relationship into which a computational output will be integrated.

Legal reform as a result of technological innovation is viewed as a question of authority, interpretation, discretion, professional obligations, participation and remedies. Faster processing times are not sufficient to demonstrate that a person has been provided with a lawful determination. Similarly, a minor adjustment made technically may significantly impact an individual's ability to obtain a hearing and/or correct errors in the record. Improvement and transformation are two distinct issues. One defines the nature of the change to the manner in which the law operates. The other evaluates the change in terms of the duty owed by the institution to provide fair treatment to individuals who have had their rights affected by the change.

Regional policy governing AI exists in the form of the African Union’s Continental Artificial Intelligence Strategy. In addition to providing a regional framework for domestic choice, it assists in identifying gaps related to regulation and amendments to existing frameworks (African Union Commission, 2024, §2.4.1).

This study is a comparative legal study that utilizes practical tools. It compares certain aspects of legal development in Pakistan, India, Kenya, South Africa, Brazil and Colombia utilizing legislation, judicial decisions, internal rules adopted by institutions and a limited set of scholarly literature. The primary contention of this book is that AI changes how laws operate in practice because AI causes a material shift in control over the classification / interpretation / correction of matters that are legally significant. The primary accountability concern is that such shifts occur without concomitant allocations of authority and mechanisms of redress. Thus, the focus of this comparative study is not simply on the technology used in each country. Rather, it focuses on the path through which institutions move from information to action and redress. Each of the countries examined presents a unique comparison and should not be taken as representative of all countries nor as examples of similar deployments.

In examining this issue, the organizing question posed by this book is what must remain subject to knowledge and contestability when legal services are being organized around a computer-based system? To answer this question requires more than merely describing the model used. Rather, it requires establishing both the source and limits

of an institution's power, the records maintained by an institution regarding persons whose rights were impacted by the use of the system, the representations made by those records, the roles assumed by legal professionals, and the avenues through which individuals may seek corrections of their rights when negatively impacted by actions taken pursuant to the system. All these factors will influence the selection of appropriate comparators. Systems that generate better records, add staff, simplify forms and create less complex alternative non-AI systems are to be evaluated alongside other assessments.

Section 1.1 sets out the study's contribution and the test for legally relevant transformation. Section 1.2 examines how customary and religious jurisdiction expose problems of digital representation. Section 1.3 develops the authority analysis through a referral problem. Section 1.4 explains the comparisons and sources assessed. Section 1.5 evaluates three possible intake designs. Chapters 2-5 extend this framework through professional obligations, data governance, evidence, liability and implementation. The instruments use the shared record introduced in the reader's guide and illustrated in Chapter 2.

1.1 AI Systems and the Meaning of Legal Transformation

A legal analysis of AI begins with a description of the work being performed. Product names, model size, and claims of intelligence are insufficient because the same application may occupy very different positions in a legal process. An internal research assistant, a public information service, and a system that prevents an application from proceeding may share an interface while presenting different questions of authority and responsibility. The classification used in this book therefore connects the system's technical function to the institutional action that follows its output.

Digitisation, automation, and inference

Digitisation concerns the form in which information is recorded and accessed. A paper file becomes a digital image; its text may then be transcribed, indexed, and searched. These operations do not, by themselves, decide what the record means or whether its contents are true. Their legal importance lies initially in fidelity, provenance, accessibility, and correction. A missing page, an incorrectly rendered name, or the loss of a qualification can affect later work even when the institution has introduced no predictive model. Digital availability and evidential reliability are different properties of the record.

Automation concerns the execution of a process with reduced manual intervention. Sending a reminder on a specified date, applying a fixed calculation, or moving a file after a recorded event can be automated without the inferential features ordinarily associated with AI. Nevertheless, the programmed rule still requires scrutiny. A remind-

er may be merely helpful, whereas treating the same date as an automatic ground for closure may affect access to the process. The absence of AI does not remove the authority question, and the presence of AI does not create it for the first time.

The OECD's explanatory memorandum describes AI through inference from inputs towards outputs that can include predictions, content, recommendations, and decisions. It recognises differences in autonomy and adaptiveness and encompasses machine-learning and knowledge-based approaches (Organisation for Economic Co-operation and Development [OECD], 2024, pp. 4-8). This is a conceptual reference, not a domestic legal definition for the six jurisdictions. It also cautions against equating AI exclusively with generative chatbots or assuming that every system using explicit rules falls outside the category.

Two descriptions are consequently needed. The technical description identifies the operations, inputs, and outputs. The institutional description identifies the purpose, user, consequence, and capacity for reconsideration. A sophisticated model restricted to suggesting search terms may exercise less practical influence than a simple rule that automatically rejects an incomplete form. Technical complexity is relevant to testing and explanation, but it is not a sufficient measure of legal consequence. Where an applicable enactment defines AI or automated processing, that definition must be examined separately from the working vocabulary adopted here.

Search and classification

The "search" function identifies and organizes existing data. While it will immediately provide a set of potential sources of information (and does so with great speed), a search function does not assess the validity of those sources. Therefore, we need to consider four separate questions. First, what is contained in your source? Second, what you have left out? Third, how do you keep current? Fourth, why did something appear?

Even after identifying sources, a researcher must determine if each identified source represents a binding piece of law or just comment on such. Furthermore, the researcher must differentiate between a current piece of law and historical reference. In addition, while improved searching can help to locate relevant case-law, it will never replace the process of analysis that makes retrieved law useful.

When the limitations imposed by searching become institutional barriers to action, searching then becomes institutionally significant. For example, assume there exists a research interface that provides only selected appellate court decisions and does not allow users to search the underlying database. Any attorney would likely rely upon it to initially orient him/her about the law, however treating its results as the complete story would be an entirely different matter. The assessment here is not simply concerned with the number of documents that were lost due to the limited nature of the

search. Rather it is concerned with the relationship among collection boundaries, user expectations and verification.

Thus, even though all of the items returned through the search are valid (i.e., they exist and represent decisions) a search conducted within a narrow collection will remain narrow.

Classification assigns subject material to categories based upon characteristics of that subject material. Classification can serve many purposes including assigning an item into a specific category or type, routing complaints, designating correspondence as urgent or routine or marking applications for additional review. The legal impact arises from the consequences associated with the assigned category. Although some labels may simply indicate where an item is stored, others may designate an item for special handling or indicate whether an employee should read the file. Categories also require a source of meaning. Therefore, an organization must differentiate between a category designation that serves as a convenient means of categorizing an item versus a category designation that represents a legitimate legal basis for treatment. An organization must also define how to handle ambiguously classified or overlap classifications.

Prediction, translation, and generation

Prediction refers to an estimation of a particular future occurrence or condition. Therefore, the target needs to be defined carefully for legal use. Determining the amount of time needed to perform a certain activity is different from determining the probability that the outcome will be favorable. Neither estimate determines what the law allows or prohibits, or whether the law is being applied correctly. Establishing agreement with past results can indicate that there is an ability to produce the reported behavior of institutions. However, this does not establish that the behavior is correct nor that it is applicable to a new situation. An estimation becomes a decision-making tool only upon an institutional decision regarding the intended use of the estimation.

The importance of establishing that distinction applies even if the predictive analysis appears statistically strong. For example, determining how likely an applicant will finish an online application form may assist in identifying an applicant's need for assistance. If using the same statistical estimate to exclude applicants who are anticipated to require greater support than others, then you are pursuing a different goal. The model alone cannot provide guidance on which goals are authorized and/or permissible. The assessment should describe the relationship between the estimated event, the applicable legal function(s), the party involved, and the consequences of incorrectly estimating an event.

Translation transforms written information into written information in another language. Translation has applications ranging from providing basic information to communicating legally binding instructions. Even though translations are designed to be

readable they can alter a classification, due date, etc., as well as change what constitutes permission versus obligation. The relevant document should include the original version of the translated document, the status given to the translated document, who reviewed the significant portions of the original document, and a method for disputing phrases in the translation. What is required in terms of certification or other forms of formal procedures depend on applicable regulations. Fluency appears to be insufficient to demonstrate compliance with such regulations.

Generative models create new versions of written words including summaries, answers, arguments, and potential rationales. Generative models differ from retrieval systems regardless of whether or not a generative model creates citations. NIST's Generative AI Profile describes confabulation as the confident presentation of erroneous or false content and notes that invented logic or citations may accompany a response (National Institute of Standards and Technology [NIST], 2024, §2.2). As a result, for legal activities there are two separate types of verification processes: first verify that a cited source exists and second verify that the source supports the rationale provided to substantiate it. These types of verification do not follow from the tone or thoroughness of the response.

Another issue occurs because a generated summary omits information instead of invents. Although an abbreviated version of a document may be free from any inaccurate statements and omit important exceptions / controversy, these omissions may affect decisions made based on the document. Therefore, verification must examine the relationship between the summary and the full documentation, not merely assess the accuracy of each statement within the summary. Likewise, generated rationales present similar issues as verification must compare a generated rationale with the actual rationale relied upon by the decision-maker. A good reason for obtaining a specific result does not equate to a recording of why that result was obtained.

Assistance, procedural control, and determination

In addition to being able to create multiple functions in a single workflow, a system can obtain guidance, translate an individual's account into another language, categorize the problem, assess the level of urgency associated with the problem, and assist in drafting a response. However, describing the last screen as simply a "chatbot" hides the fact that there are multiple separate steps involved in this interaction and that there is potential for error in each step. Each function will have to be corrected individually if the retrieval of guidance was incomplete; translation lost some meaning; categorized incorrectly; or if the wording used was not supported by previous statements. Therefore, the legally relevant system consists of the links between these interactions, the interface (i.e., personnel) where they will be encountered and the ultimate institutional actions taken as a result of those interactions.

Therefore, assistance and determination represent relationships in a process, not permanent characteristics of a product. Assistance provides materials available for review

and assessment independently. Procedural control determines how a matter is received into, moves through or is released from an institution. Determination establishes or has the effect of establishing the nature of the institutional response. A system that is designated as advisory can determine the institutional response if its output is effectively required. Likewise, while an automated computation can provide one of many components that together establish an institution's decision, it can do so in such a way as to make it clear what parts of the decision were made on behalf of whom.

More than just providing a signature is necessary for meaningful human control over the operation of an automated system. In order to evaluate the output produced by a reviewer, he/she/they need access to the underlying data used to produce the output, he/she/they should possess relevant competence, he/she/they should have adequate time to verify the accuracy of the output, and he/she/they should have the authority to deviate from the recommended course of action. Furthermore, disagreement regarding the recommendation provided must be capable of resulting in a different institutional action rather than merely annotating the recommendation without changing its outcome. This represents a condition that should be examined in terms of the operational workflow of an institution. While a low override rate may indicate that the reviewer has been provided reliable support by an automated system, it may equally suggest that there has been insufficient oversight / monitoring by reviewers, and/or that an environment exists that discourages deviation from the recommendation provided by an automated system. It is therefore important to consider both the actual work flow documents and the working conditions under which reviewers operate to differentiate among these alternatives.

Human control must be examined when the person first experiences a consequential output. A later appeal may reverse an erroneous decision without restoring a missed deadline or lost opportunity. Reserving the final decision for a human does not remove burdens already imposed through a document request, routing rule or priority assignment. Analysis must therefore examine entry requirements, requests for information, notices and opportunities for correction as well as the final determination.

The threshold for legal transformation

The distinction between the formal transformation of law (the official establishment of new legal structures) and operational transformation (how existing authority is exercised differently), highlights the differences in the way formal legal transformations occur versus operational transformations. Formal transformations establish new powers, duties, procedures or enforceable rights by recognized legal processes. Operational transformations involve the material alteration of how authority is practiced - such as encountering different sources, structuring discretion differently, verifying professional activities performed differently, or establishing and challenging claims differently. While there may be some overlap between formal and operational transformation, both types of transformation should not simply be presumed solely based upon installing software.

There exists a type of interpretative change that arises when a computational arrangement performs more than just locating law, and instead provides for a specific interpretation of it. For example, a decision support system's use of an open-ended standard could be presented as a menu of closed factors. The key issue would be determining whether the menu assists in evaluating the considerations, or whether it excludes other considerations related to law. Discretion may still exist -- it will have merely moved from case level judgment to judgments made at the time of design. By identifying who approved the design choices, and the process for revising them, the transition will be analyzable.

A professional change involves the reallocation of resources (research, drafting, supervision, verification etc.). If the production of a first draft is completed quickly, then the same amount of time previously allocated to analyzing the details of a draft may now be utilized for more detailed analysis. Conversely, if less time is needed for creating drafts, the "time saved" may be absorbed into additional tasks. Neither circumstance is assumed here. The concern is whether the allocation of responsibilities for performing the various tasks creates a change in how professionals exercise their responsibility or the context in which they exercise their professional judgment. For example, an organization that assigns fewer tasks to individuals, yet fails to assign someone to verify that the assigned tasks were properly executed has created a greater degree of transformation than simply changing the order in which tasks are performed.

Remedial change is concerned with issues of discovery, explanation, liability and remedies. A party seeking redress must identify which document(s) were used to produce their case, which entity(s) rely(ed) on those documents, and who is capable of changing the outcome. An application that facilitates access to these three steps can enhance accountability without eliminating the decision-making entity. Conversely, an application that disperses these steps may hinder a party's ability to seek redress. The determination of remedial change tracks redress through correction and/or relief -- and not solely the initial reaction of the responding entity.

If an overly broad definition is employed, virtually all administrative modifications would qualify as a form of transformation. Establishing a materiality requirement precludes this result. The analyst must specifically identify a tangible mechanism, a legally relevant relationship, and sufficient evidence demonstrating a significant impact resulting from the change. Although retrieving information more rapidly without altering the authorities responsible for providing that information or verifying it effectively represents modernization; however, modifying an interface so that it systematically eliminates routes to redress may represent a form of transformation. Classification as such is purely descriptive -- whether each type of modification is justified remains a separate evaluative criteria.

Contribution to Debates About Automated Legal Power

Scholarship on three fronts frames the argument advanced in this book. Citron's examination of how automated administration functions in the U.S. relates procedural protections to how systems implement public policies and explain individualized decisions; she advocates using audit trails and limitations on uncritical reliance on computer-generated results (Citron, 2008, pp. 1305-1306). This demonstrates why design technology should fall under the umbrella of legal process. However, due-process doctrine as it applies in the U.S. will not apply domestically as articulated in this book. Instead this book takes the institutional aspect of authority, procedure and remedy into domestic jurisdictional contexts.

According to Selbst and colleagues (2019, §2.1), fair analysis of models fails when models are abstracted from the social arrangements through which their outputs gain meaning. This implication accepted by this book suggests that legal validity cannot be determined through a model-based performance metric. A routing score and a refusal to file are distinct objects of evaluation despite one producing the other. The book transforms this insight into an inquiry regarding the actor entitled to assign a procedural consequence to an output.

Cobbe, Lee & Singh recognize reviewability throughout the decision making process (commissioning, developing, operating and investigating), while also recognizing that over-documentation can obscure scrutiny; pose privacy risks; and fail to address systemic inequalities of power (Cobbe, Lee & Singh, 2021, sections 3-5). The current book expands on this position. It does not argue that records; explanations; or human reviewers are novel mechanisms. Rather, it connects a limited record to a particular legal dispute; an authorized reviewing actor; and a potential corrective action.

In doing so, Cobbe, Lee & Singh's approach leaves a practical dilemma that becomes apparent in the comparative analysis provided in this book. Documenting data more comprehensively may facilitate an external reviewer's understanding; yet documenting data more extensively may impose costs that smaller institutions cannot afford. Common digital categories may facilitate exchanging data among parties; yet common categories may misrepresent legitimate institutions recognized by law. Similarly, formally

assigning human review may conceal the loss of discretion if officials routinely accept classifications produced by suppliers. Therefore, we cannot conclude that everything should be documented in every instance.

The contribution is thus an integrative and comparative one. The book specifies a materiality test for operational transformation, explains how legal recognition and digital representation can diverge, and links safeguards to the authority and resources needed to make them effective. Its instruments express these propositions as decisions that can be examined and contested. They remain normative proposals where domestic law does not itself prescribe them. The claim would be weakened by evidence that the relevant legal consequences remain unchanged despite the hypothesised relocation of control, or that a less burdensome arrangement supplies equally effective correction.

Applying the materiality threshold

Consider three constructed versions of the same legal-aid intake service. In version A, a classifier places applications in suggested folders. Every application is read in the ordinary queue, original narratives remain available and staff can correct a label without affecting eligibility or time limits. Improved speed alone supports a finding of administrative modernisation. To call this transformation would require evidence of a further consequential change, such as altered responsibility for assessing claims or systematic exclusion from professional attention.

In version B, the identical classifier marks missing documents and closes the file unless an applicant uploads them within seven days. Assume for this example that the governing procedure allows applicants to explain unavailable documents and does not authorise that closure rule. The change now reallocates practical control over access: a model output triggers a procedural consequence, the applicant bears the burden of reopening, and an exception may never reach an authorised officer. The mechanism and legal relationship are identifiable even if average classification accuracy is high. This is operational transformation; on the stated assumptions it is also legally defective.

Version C is the borderline case. Staff must review every flag before any consequence, but the institution doubles their caseload after introducing the system. Formal responsibility remains unchanged. A transformation finding needs evidence that reduced review time materially changes consideration of exceptions, handling of disagreement or effective access. Relevant evidence includes the time available, sampled original files, reasons for overridden and accepted flags, complaints and resulting corrections. A low override rate alone is inconclusive. If adequate review survives the workload change and affected persons face no material alteration, the more defensible classification remains modernisation.

Materiality is established when the alteration is capable of changing a legally relevant opportunity or outcome and does so through a demonstrated institutional mechanism. A minor isolated interface defect is an error; a configured rule that repeatedly fore-

closes an available route changes the operation of the procedure. A single grave exclusion can nevertheless be material for the affected person. This test does not require a universal numerical threshold. It requires an explicit comparator, an identified consequence, evidence at the point where that consequence arises, and consideration of an explanation under which no transformation occurred.

A function-based classification record

Table 1.1 brings together the technical and institutional descriptions. Its entries state possible consequences and assessment questions, not uniform legal requirements or findings about particular deployments. The same record should also describe a feasible non-AI alternative. This keeps the analysis directed towards the legal task rather than towards proving the value of a preferred technology.

Table 1.1. System functions and possible legal consequences

Function	Immediate output	Possible legal consequence	Assessment focus
Digitisation and transcription	Digital image or searchable text	An omission or transcription error changes the material available for consideration.	Original retained; fidelity, completeness, provenance, access, and correction checked.
Rule-based automation	Calculation, reminder, or workflow action	A programmed condition affects time limits, routing, or continuation of a matter.	Rule traced to its source; inputs, exceptions, and failure consequences identified.
Document search	Retrieved or ranked sources	Collection and ranking shape which authorities the user encounters.	Coverage, legal status, currency, jurisdiction, and access to the full source examined.
Classification	Category, flag, or queue	A matter receives different scrutiny, procedure, assistance, or access.	Meaning and approval of categories; effect of each label; exceptions and rerouting.
Prediction	Estimate, score, or forecast	Recorded patterns become influential criteria for present action.	Target meaning, lawful relevance, validation population, uncertainty, and error consequences.

Table 1.1. System functions and possible legal consequences

Function	Immediate output	Possible legal consequence	Assessment focus
Translation	Text in another language	A person's account, obligation, or procedural instruction changes meaning.	Original and operative versions; material wording; checking and dispute route.
Generation and summarisation	Draft, answer, or compressed record	Unsupported statements or omissions enter advice, evidence, or reasons.	Factual and legal verification; completeness; source provenance; accountable approval.
Determination or execution	Operative response or action	An output changes a person's position without substantive intervening judgment.	Specific authority; decision criteria; participation, review, correction, and relief.

The operational owner, legal reviewer, and technical team should complete the record together, identifying the version, intended users, sources, connected functions, and point of consequence. A change from internal assistance to public communication, procedural gatekeeping, or outcome determination requires renewed assessment. The resulting record establishes what is being adopted and why the differences between functions matter. It does not determine, without the jurisdiction-specific inquiry that follows, whether that use is authorised.

1.2 Legal Pluralism and Institutional Variation

A legal system cannot be represented adequately by the materials that happen to be available in a searchable database. Nor can an institution be understood solely through its formal title. The conditions of AI adoption depend on how authority is distributed, how people move between procedures, how accounts become records, and which institution can recognise or correct a legal position. Legal pluralism directs attention to these relationships without assuming that every socially influential practice possesses the same status in state law.

The Global South is used here as a comparative orientation towards unequal conditions of legal knowledge, institutional capacity and control over technical infrastructure. It is not a legal category or an explanation in itself. South Africa and Kenya expose different forms of constitutional recognition within plural legal orders; Colombia and Brazil provide different judicial governance instruments; Pakistan and India supply distinct evidence and administrative questions. The comparisons examine how these differences change the conditions for lawful assistance and effective challenge.

They do not attribute every difficulty to colonial history or assume that all selected settings face it equally.

Legal recognition, social authority, and practical access

Legal recognition, social authority, practical access and digital visibility are distinct. Recognition concerns the status and powers an arrangement has under applicable law. Social authority concerns the influence it exercises in a community. Practical access concerns whether a person can use or leave the process and obtain assistance or review. Digital visibility concerns what the selected records and interface reveal. None establishes the others automatically. The South African and Kenyan examples below show why a system must preserve these distinctions rather than treat every searchable institution or practice as legally interchangeable.

Living customary law and the limits of a historical corpus

South Africa offers a focal example of a legal order in which recognised law cannot be reduced to a fixed historical collection. In Shilubana v Nwamitwa, the Constitutional Court considered the development of customary law concerning traditional leadership. Its reasoning emphasised both historical practice and the community's capacity to develop its law, within constitutional constraints. The Court warned against treating colonial records as an unproblematic statement of customary law and required attention to contemporary practice, rights and the circumstances of change (Shilubana, 2008, paras. 42-49, 75; Constitution of South Africa, ss. 39(2), 211).

Now consider a constructed research assistant trained only on older reports of succession practices. It returns an apparently categorical male-only rule and ranks later community developments as inconsistent exceptions. The system has not merely missed a document. Its representation mistakes historical evidence about law for an exhaustive account of the law's present sources and development. The legal defect would persist if every quotation were perfectly reproduced. The relevant corrective design preserves the date and institutional source of each proposition, retrieves subsequent legal development, and refers unresolved customary-law questions for a properly supported legal assessment.

Shilubana does not authorise a model to infer current community law from a database or entitle a self-appointed spokesperson to settle it. The case supplies a method for treating history, present practice and constitutional norms as legally relevant; evidence in a new dispute must still be established. Where community accounts conflict, the record should preserve that disagreement rather than encode one account as ground truth. The scholarly point is that inclusion cannot be measured simply by increasing the number of customary records. A larger corpus can reinforce a mistaken source hierarchy if its retrieval and classification rules freeze a living legal order.

Religious jurisdiction and the consequences of digital referral

Kenya supplies a different constitutional structure. Article 159(2)(c) recognises alternative dispute resolution, while article 159(3) limits traditional mechanisms where they conflict with the Bill of Rights, justice and morality, the Constitution or written law. Article 170(5) confines Kadhis' courts to questions of Muslim law relating to personal status, marriage, divorce or inheritance in proceedings where all parties profess Islam and submit to that jurisdiction (Constitution of Kenya, 2010, arts. 159, 170). Recognition therefore operates through specified conditions; it is neither unlimited nor interchangeable across institutions.

Suppose a public information service routes an inheritance query to a Kadhi after inferring religion from a surname. The inference cannot establish all parties' religious affiliation or submission to jurisdiction. Even accurate prediction of one applicant's religion would leave the other conditions unanswered. A lawful information design can explain that the route may be relevant and identify its conditions. A system that treats the prediction as conclusive jurisdictional assignment would substitute an unproved proxy for constitutionally material facts. Whether a subsequent filing is within jurisdiction remains a question for the competent legal process.

The comparison reveals two different representation errors. The South African example freezes a source of law that may develop; the Kenyan example converts conditional jurisdiction into a demographic category. Both demonstrate how standardisation can alter relationships among institutions without formally changing their powers. They require different remedies: source-sensitive treatment of legal development in the first, and explicit jurisdictional inquiry with effective correction of routing in the second. Neither is resolved simply by improving a language model's aggregate accuracy. These examples give the Global South framing analytical content while producing questions that can also travel to other plural legal orders.

From a person's account to a usable record

Beginning representation occurs prior to digitization. An individual provides a description of a problem; an intermediary determines which elements will be documented; forms provide the structural categories used to document the problem; procedures determine the relevant evidence needed to address the problem; publications determine which items from the documentation process will be disseminated to a larger audience. An AI system using published legal materials receives the account preserved in those materials; an intake system may instead receive an earlier version. The difference between the original account and the institutional representation of that account is critical. There is no reason to assume that a model can recover a qualification, ambiguity or alternative explanation present in the original account but absent from the input provided to the model.

An illustrative example of potential differences based on the format of intake illustrates how this can occur. In the first instance, an applicant submits a description of a dispute concerning housing (accommodation), employment, and family responsibility. The intake form permits submission of only one primary category. As a result, the description submitted by the applicant is classified as a "housing" issue. A subsequent classification algorithm uses the housing field as part of its training set and directs other descriptions classified similarly into a housing queue. While there are many reasons that may contribute to the resultant pattern (e.g., model error) it may have begun with the rules governing the initial recording of the applicant's account and become increasingly important as the category determines which advisor(s) or procedure(s) receive the account.

Therefore, the provenance of a field includes more than simply where it exists within a database. It includes who originally furnished the information, who reformulated it, why a particular category was selected and whether it represents an allegation, an observation, an expert opinion or a conclusion reached at the end of an investigation. Any representations made regarding these characteristics should persist throughout any reuse. For example, a "resolution" entered in a field may signify agreement, withdrawal, referral or administrative closure. Unless defined, the apparent resolution presents an ambiguous foundation for either predicting or evaluating.

Similarly, absence can have multiple interpretations. A practice may not produce documentary records; documents may exist but remain un-digitized; digital documents may be inaccessible; accessible documents may reside outside the scope of the application's repository; or the retrieval / interpretation processes used by the system may fail to identify / access them. These are distinct types of gaps. Depending upon the nature of the gap, an appropriate response might include improved indexing; arrangements facilitating access; narrowing claims regarding coverage; or referring requests to individuals possessing relevant expertise. Treating each type of gap as an opportunity to obtain additional data ignores both legal constraints and the fact that the proposed task may exceed the limits of available evidence.

Language and the whole procedural journey

Language access has to be examined across the complete process. A person may be able to enter a narrative in one language while the categories, requests for evidence, generated reasons, and challenge procedure operate in another. A multilingual interface does not establish that the person can check how the institution has understood the account. Translation quality is therefore related to participation: an error matters differently when it appears in an internal search query, a recorded statement, or a notice whose meaning affects the next procedural step.

The AU Strategy expressly identifies linguistic diversity, gender, disability, and underserved communities as concerns for inclusive AI development (African Union Commission, 2024, §2.4.4.1). That is a policy basis for asking distributional questions, not

evidence of accessibility in a particular application. The operational inquiry remains specific. It asks which language is used at each stage, which version is authoritative, who can correct it, and whether assistance is available without depriving the person of control over the account.

Rural access similarly concerns a journey rather than a connection. An online entry point may reduce the need for an initial visit while leaving verification, document collection, or review dependent on travel. The relevant comparison includes distance, expense, waiting, assistance, and the possibility of completing the task through another channel. These are matters to measure in the selected institution. Their importance cannot be inferred from a general description of rural life, and an improvement at one stage cannot establish a net improvement across the whole process.

Gender, disability, and documentation status require the same attention to mechanisms. Who controls the device receiving a notice? Can a person communicate privately? Does the interface work with the necessary assistance or accessibility arrangements? Can someone lacking a particular identifier seek advice before the institution determines whether that identifier is legally required? These questions do not assume uniform disadvantage or identical experiences. They identify conditions that can make a formally available route unusable and that would remain invisible if the evaluation counted only completed digital submissions.

Institutional influence through selection and routing

The relationship between institutions may also shift as a result of using an AI Application for creating and maintaining relationships. This is due to the ability of the application to manage visibility, categorization, and referrals. When a user selects which data will be made searchable, this has implications in terms of what users (e.g., other employees) have ready access to. As the developer creates categories, these may affect where items are placed. And, when the user controls corrections, there may be a dependency created if an error occurs that crosses organizational boundaries.

These are tools with which to investigate, but not conclusions regarding whether AI either centralizes or decentralizes authority. For example, a single service could enhance collaboration among entities; alternatively, it could render local activity dependent upon decisions that local officials cannot see or modify.

A hypothetical referral application linking a public agency, a legal aid organization, and a community mediation organization presents the user with a menu representing a particular perspective on possible pathways. If a specific service is omitted from the menu, users relying on that menu may not discover it. An automated referral may mask assumptions related to suitability, consent, or the type of conflict at issue. While the referral platform establishes a pathway for individuals, it does not provide jurisdiction over them simply because it sends them somewhere. The platform's institutional assessment should determine who established the routing rule, what options are avail-

able to the destination entity, and how referrals that were sent inappropriately can be re-examined.

Although private vendors are represented on the same map as governmental bodies, private vendors' functional roles may involve making data selection decisions, updating category names and descriptions, modifying interfaces or providing the sole method of obtaining historical information associated with outputs generated by applications. There needs to be clear separation between technical management and legitimate authority. Public institutions' responsibilities cannot be determined solely based on their organizational charts given critical knowledge or remedial capabilities reside outside of those organizations. In addition, while private vendors' technical significance may indicate they possess significant capability, this does not mean they exercise public power via the use of the systems.

Recognition without idealisation

Attention to plural arrangements does not require treating every practice as legitimate or every local representative as the voice of a community. Internal disagreement, coercion, unequal participation, and the ability to leave a process are part of the institutional inquiry. A practice can be socially influential and still conflict with applicable rights. Its recognition as an object of study is not an endorsement of its outcomes. The same scrutiny applies to formal institutions: official status is not evidence that a procedure is accessible or that every participant's account is adequately represented.

The objection that an institutional map could become unmanageably extensive is important. An application cannot incorporate every normative tradition or every possible legal experience. The appropriate response is a defined scope and an intelligible route beyond it. A service confined to published procedure can be useful if that boundary is clear. It becomes misleading when its restricted collection is presented as a complete account of law, available institutions, or possible remedies. Honest limitation is more useful than apparent comprehensiveness built from unexamined omissions.

Standardisation need not eliminate institutional difference. A common record format can preserve the source, status, uncertainty, and limits of an entry. A uniform substantive category may instead remove distinctions that matter to the receiving institution. The analysis should identify precisely what is being standardised and why. Interoperability concerns the ability to exchange usable information; it does not, by itself, justify treating every institution's categories or every person's circumstances as equivalent.

An institutional authority and knowledge map

Table 1.2 connects the functional record to the institutions that supply knowledge, receive referrals and correct mistakes. Read it against the South African and Kenyan analyses above: the map must preserve differences in legal recognition, contemporary practice and conditional jurisdiction.

Table 1.2. Institutional authority and knowledge map

Actor or position	Authority and knowledge to establish	Representation or relationship to examine	Record and corrective route
Court or tribunal	Jurisdiction, procedural control, review powers; pleadings, evidence, orders, and reasons.	Which records enter the system, and what legal status is assigned to each?	Pinpoint authority; collection limits; record custodian; competent reviewing office.
Ministry, regulator, or registry	Assigned function; application and decision powers; registers and administrative records.	Does a category alter scrutiny, eligibility, priority, or access to consideration?	Source mandate; category definitions; decision owner; reconsideration and record-correction route.
Frontline staff and intermediaries	Intake role and practical knowledge of recurring problems, documents, and exceptions.	Which parts of a person's account are retained, reformulated, or excluded?	Intake rules; original account where appropriate; exception records; escalation authority.
Customary, religious, or community arrangement	Specific recognition, subject matter, participation, enforcement, and review; local knowledge.	Whose account represents the arrangement? Can a participant disagree, refuse, or use another route?	Domestic recognition sources; referral and consent conditions; internal disagreement; competent review.
Lawyer or legal-aid organisation	Professional and organisational responsibilities; client narratives, advice, and procedural knowledge.	Does the system improve assistance or create new restrictions or dependence?	Applicable duties; supervision and verification; referral criteria; complaint and remedy channels.
Supplier, operator, or data custodian	Contractual functions; access, hosting, updates, logs, and control of source records.	Does practical control exceed the institution's ability to understand or supervise changes?	Responsibility and version records; evidence access; source correction; suspension and continuity.

Table 1.2. Institutional authority and knowledge map

Actor or position	Authority and knowledge to establish	Representation or relationship to examine	Record and corrective route
Affected people and representative groups	Participation and challenge routes; first-person knowledge of circumstances, access, and error.	Who cannot enter, understand, correct, or contest? Which experiences are absent from completed files?	Accessible input and assistance; coverage limitations; complaint record; route beyond system scope.

Legal and operational staff should prepare the map with records specialists and, where appropriate and ethically feasible, affected-person input. Each entry needs a supporting source, a responsible reviewer, and an indication of uncertainty. A disagreement about recognition or referral should remain visible rather than disappear into a single consensus label. The completed map establishes whose authority and knowledge the application operationalises, what remains outside it, and where a person can take a matter that the system cannot adequately represent.

1.3 Legal Authority for Adopting and Using AI

Authority to adopt AI is not a single permission. Buying software, disclosing a file, creating an inference, changing a procedure, approving a communication, and issuing a decision are different acts. Their legal bases may come from different instruments and involve different competent bodies. A general approval for innovation cannot answer all of these questions. The analysis must start with the proposed conduct and follow it through the institution's powers, the conditions attached to those powers, and the remedies available when their exercise is challenged.

Authority analysis proceeds act by act. The Kenyan referral problem below shows how to distinguish administrative assistance from a substitution for legally required judgment. The comparison with judicial AI rules in Chapter 2 and the South African administrative-law analysis in Chapter 4 develops the same inquiry through different domestic sources.

The underlying function and the competent institution

The starting point is the legal function that the system is intended to support. "AI in courts" is too broad to establish a mandate. Preparing an internal summary of public judgments, checking the format of a filing, and refusing to accept a filing are different activities. Each description should identify the institution, responsible office, affected population, relevant territory, and contemplated consequence. Precision at this stage prevents a narrow approval from being treated later as permission for a more consequential use.

For a public body, the authority file begins with the instrument assigning the underlying function. Depending on the jurisdiction and institution, the relevant sources may include constitutional provisions, legislation, delegated instruments, court rules, and valid internal authorisations. The exact provision matters. A statement of institutional objectives is not necessarily a grant of every power that might help achieve them. Where an incidental or implied power is relied upon, the argument must explain its relationship to the assigned function and address any express limitation or contrary authority.

Competence also has an organisational and territorial dimension. The office entitled to approve expenditure may not be the office entitled to amend procedure. A central body may provide technical services while another institution retains the substantive decision. The question is not whether senior officials support the project but whether the relevant act falls within the approving actor's authority. A committee can coordinate the assessment, yet its existence does not settle which member or office possesses the legally necessary power.

Constitutional, administrative, and sectoral constraints

A constitutional inquiry connects the proposed conduct to actual provisions and authoritative interpretations. Institutional independence, equality, privacy, participation, and access to justice may be relevant, but their applicability and legal tests cannot be assumed to be identical across jurisdictions. The analysis must identify the person whose interest is affected, the body responsible for the conduct, the nature of the interference, and the route through which it can be examined. A list of rights without those connections does not establish how a system may lawfully operate.

Administrative-law analysis concerns the exercise of an existing power as well as its existence. A proposal may raise questions about relevant considerations, the treatment of exceptions, adequate inquiry, notice, reasons, and reconsideration. Whether any particular requirement applies must be established from the governing law. A technically reliable model may nevertheless use a consideration that the institution is not entitled to make decisive. Conversely, disagreement with a model output does not establish unlawfulness unless the relevant factual and legal conditions are examined.

Sector legislation and procedural rules can be decisive because the software serves a particular legal task. A registry, licensing office, court, legal-aid service, and regulated professional practice do not become legally interchangeable through their use of AI. The applicable sources may specify documents, time limits, decision criteria, access restrictions, or opportunities to respond. A proposed system must be assessed against those requirements. General technology policy can guide implementation, but it cannot be treated as a complete account of the legal framework governing the service.

The same discipline applies to the form of a necessary change. An internal instruction may be sufficient for a permitted working arrangement, whereas a procedural amend-

ment may require action by a rule-making body. A condition fixed by legislation may require a different legal route again. The authority record should identify the source of the existing condition and the actor able to change it. Describing a modification as a configuration update does not settle its legal character.

Purchase, data use, and procedural change

Procurement authority concerns obtaining goods or services, committing funds, and entering an agreement. The record should identify the applicable purchasing framework, budgetary approval, signatory, and contractual scope. Those matters are necessary to understand the acquisition, but they do not by themselves establish the lawfulness of every use of the acquired system. A properly authorised contract may still describe an application whose data access, procedural effect, or decision role requires a separate legal basis.

Data activity needs to be disaggregated with equal care. Collection, inspection, disclosure to a supplier, retention, model development, validation, and processing of a current matter are not one event. An institution's ability to hold or read a document does not establish unrestricted permission to give it to another operator or permit secondary use. The analysis must identify the source, purpose, recipients, access arrangements, and proposed onward uses, then connect each to the applicable legal and professional framework.

Public availability also requires a precise meaning. A document may be accessible for one purpose or through one channel while its aggregation or reuse raises additional questions. Confidentiality, restrictions on court records, contractual conditions, and sectoral duties may need separate examination. The authority assessment should not treat the removal of a name as proof that all relevant restrictions have disappeared. It should state what information is actually processed and why the proposed arrangement is lawful, rather than allowing a technical description to substitute for that conclusion.

Procedural changes concern what officials and affected people must do. A new required field can impose a documentary condition. Automatic routing can change the body considering the matter. A generated notice can affect how a person understands the next step. These effects should be compared with the existing procedure, identifying which are permitted operational choices and which need a different authorisation. The relevant distinction is between implementing a lawful requirement and introducing a new requirement through the software.

Assistance, delegation, and the exercise of judgment

Delegation analysis begins by identifying the function entrusted to an actor, not by asking whether a computer has been used. Locating a source, calculating an amount, drafting a paragraph, assessing evidence, and exercising a discretion are analytically distinct. The governing law determines what assistance is permissible, what can be performed through another person or process, and what must remain with the assigned decision-maker. The system's label cannot supply that determination.

Three arrangements should be distinguished. In the first, the system performs a supporting task while the authorised actor independently assesses the relevant material. In the second, it supplies a recommendation that structures the decision and may narrow the choices practically considered. In the third, it produces an operative result without meaningful case-level judgment by the actor to whom the function is assigned. The distinction directs legal inquiry; it does not declare the first arrangement lawful or the third unlawful in every jurisdiction.

Substantive control must be examined within the working conditions. An official may retain a formal power to reject an output but lack the underlying record, time, or technical access required to exercise it. A signature in those conditions does not establish independent consideration. Conversely, extensive assistance can remain compatible with retained judgment where the reviewer can assess the evidence, identify uncertainty, depart from the recommendation, and give reasons. The legal significance of these facts depends on the particular function and applicable sources.

The hypothetical legal-aid classifier illustrates the separate powers involved. Identifying a possibly missing attachment is one act. Requiring that attachment, deciding whether an exception applies, imposing a response deadline, and closing the application are others. Even where one body performs all of them, the authority analysis should distinguish their legal bases. A system configured for the first act should not silently acquire the power to perform the last. The person able to reverse an erroneous closure also needs to be identified before the workflow is approved.

Responsibility follows alongside, but does not collapse into, delegation. The institution needs to know who approves categories, verifies outputs, supervises use, investigates failures, and repairs affected cases. A contract can allocate work and liability between its parties without establishing every right available to an affected person outside the contract. The assessment should therefore distinguish internal responsibility, supplier obligations, professional accountability, and the public or private remedies applicable to the underlying act.

Private organisations and professional duties

Private adoption requires a related but distinct inquiry. A law firm buying drafting assistance, a non-profit providing legal information, and a supplier operating part of a public service have different relationships with their users. Corporate approval and contract may establish who can acquire the service, while statutory or professional rules govern how it is used. Neither private ownership nor public-facing activity alone determines the full set of obligations. The analysis follows the particular actor, function, and affected relationship.

Professional duties should be connected to retained tasks rather than described abstractly. The authority file needs to identify who checks legal sources, verifies facts, supervises staff, protects confidential information, and communicates limitations. The precise content of these duties must come from the relevant professional framework. A supplier's promise that an application is suitable for legal work does not establish that a professional can satisfy those duties under the proposed conditions of use.

The distinction between a draft and an adopted communication is particularly important. Once an organisation approves generated text for a client, applicant, or court, the inquiry concerns the organisation's act as well as the tool's contribution. The origin of the wording does not establish that its statements are correct or that responsibility lies elsewhere. Whether disclosure of AI assistance is required remains a jurisdiction-specific question. Verification and accurate attribution of institutional responsibility are separate issues that should not depend upon an unsupported assumption about disclosure rules.

Existing powers, legal change, and implementation limits

Two shortcuts should be avoided: requiring a new statute for every AI-assisted task and assuming that existing powers authorise every application. The AU Strategy's discussion of existing frameworks and regulatory gaps illustrates a policy approach that does not reduce governance to a single legal instrument (African Union Commission, 2024, §2.4.1). For the book's domestic analysis, the relevant question is the specific gap, if any, between the proposed conduct and the powers and conditions already applicable to it.

Existing authority may suffice where the application remains within the assigned function, preserves the legally required judgment, and complies with procedural and information constraints. This is a conditional analytical proposition, not a conclusion about any identified deployment. An express prohibition, an exclusive allocation of power, or a material procedural change outside existing competence may instead require redesign or action through a different legal instrument. The conclusion should explain that requirement rather than rely on novelty as the reason for legislation.

Not every implementation problem is a missing legal power. An institution may lack reliable records, trained reviewers, operational funds, or access to supplier-held evid-

ence. A new enactment would not, by itself, demonstrate that these capacities exist. The assessment should distinguish absence of authority, uncertainty about authority, failure to comply with an identified rule, and inability to operate necessary safeguards. These different findings call for different responses and should not be compressed into a general statement that regulation is inadequate.

Proportionality addresses the burden of review without removing the review itself. Internal search of public materials warrants a different depth of inquiry from an application affecting access to a process. A limited pilot may help establish feasibility, but its legal significance depends on actual conditions. Testing with constructed records differs from processing confidential files, and shadow outputs differ from outputs that influence a current decision. "Pilot" describes the arrangement; it is not evidence of an exemption.

A worked authority argument for an inheritance referral

Return to the constructed Kenyan inheritance service. The challenged act is not the purchase of a classifier; it is treating its assignment as a conclusive direction that the matter belongs before a Kadhi. Article 170(5) supplies the controlling constitutional conditions. A task limited to displaying information about possible forums leaves those conditions for determination through the competent process. A task that refuses another route solely because a surname model predicts Islam cannot establish those conditions and therefore cannot justify that refusal on article 170(5).

The strongest counterargument is that triage is administrative assistance and that a later officer can correct an erroneous referral. That argument succeeds only to the extent the arrangement preserves a real opportunity to reach that officer before an adverse consequence. The label attached to the software is immaterial if its operation prevents a filing, conceals alternatives or treats inferred affiliation as established. Conversely, the mere presence of an automated suggestion does not make an otherwise accessible information service an unlawful delegation of adjudication.

The reserved judgment is the application of jurisdictional conditions to the matter, including the constitutionally required submission of the parties; the permissible assistance in this constructed design is explaining and organising information relevant to that inquiry. The project record should identify the official empowered to change the referral and the applicable procedural route for a disputed jurisdictional decision. Kenya's Constitution places subordinate courts and bodies exercising judicial or quasi-judicial functions within the High Court's supervisory jurisdiction, subject to article 165(6); it does not make every mistaken information message a completed judicial determination. The route and relief therefore depend on what act actually occurred.

The reasoned conclusion is to permit a bounded informational suggestion, conditional on verified institutional authority for the service, while rejecting demographic inference as a substitute for article 170(5)'s jurisdictional conditions. If the operational ob-

jective is to assign proceedings conclusively, the institution must identify a lawful decision process with evidence of those conditions. A procurement contract cannot supply the missing constitutional basis. This conclusion is narrower, and more useful, than either a universal prohibition on judicial AI or an assumption that general modernisation powers authorise every automated consequence.

The legal-authority assessment record

Table 1.3 separates the acts requiring approval and the evidence supporting each conclusion. It should be completed before a live pilot and revisited when purpose, data, users, institutional role, or consequences materially change. The form does not confer authority. Its purpose is to make the underlying legal reasoning, supporting provisions, and unresolved questions accessible to the competent approving and reviewing bodies.

Table 1.3. Legal-authority assessment form

Assessment field	Evidence to enter	Conclusion or question to record
Proposal and boundaries	Institution, system and version, precise task, users, population, territory, and excluded uses.	What acts are proposed, and where can an output affect a legal position?
Underlying function	Enabling instrument, exact provision, applicable version, assigned office, and relevant interpretations.	Is the function and competent actor established? What limits remain?
Procurement and finance	Purchasing framework, budget approval, authorised signatory, contract scope, and conditions.	Who may acquire the service, and what does that approval not resolve?
Data activity	Source records, purpose, collection, testing, training, disclosure, recipients, retention, and reuse.	Separate basis for each activity; any restriction, condition, or unresolved permission.
Procedure	Existing and proposed entry, document, routing, notice, response, and closure requirements.	Operational arrangement, rule amendment, or another instrument required?
Decision and delegation	Legal allocation of the function; permissible assistance; delegation sources; actual workflow.	Who exercises judgment? Is any reserved function transferred or practically displaced?

Table 1.3. Legal-authority assessment form

Assessment field	Evidence to enter	Conclusion or question to record
Rights and professional duties	Applicable provisions, duties, scope, exceptions, authoritative interpretations, and responsible persons.	Which conditions must the use satisfy, and can the retained duties be performed?
Substantive human control	Record access, competence, time, ability to depart, authority to act, and recorded reasons.	Independent assessment demonstrated, restricted, or not established?
Challenge and correction	Competent forum; notice and information; source and decision records; available corrective powers.	Can the person contest the relevant act and obtain effective reconsideration?
Pilot and material change	Trial data and output limits; approvals; purpose, population, version, and role changes; stop authority.	What is authorised for the trial, and what requires renewed assessment?
Reasoned conclusion	Reviewers, source pinpoints, date, objections, conditions, unresolved questions, and review trigger.	Permitted scope, conditional use, clarification, redesign, deferral, or rejection, with reasons.

The final record should state the permitted scope, conditions, responsible actors, and events requiring reassessment. An unresolved essential authority is not a positive finding merely because no prohibition has been located. Equally, uncertainty should not be presented as a definitive prohibition. Clarification, restricted use, a different procedure, or legal change may be needed. The decision belongs to the competent institution acting under the applicable law; the assessment makes the basis of that decision explicit and preserves it for later scrutiny.

1.4 Comparative Research Design and Source Assessment

The research is a selective comparative doctrinal and normative study. It reads primary legal authorities for the conditions under which computational assistance may affect legal work, and uses documented institutional material to identify the difference between a prescribed safeguard and a demonstrated outcome. It does not report interviews, a deployment census or an estimate of AI's causal effects. Its comparisons are completed legal analyses of selected problems, supplemented by explicitly constructed applications that test the implications of the rules.

The unit of comparison is a legal relationship: an institution or professional, a computationally assisted task, an affected interest, the authority attaching consequences to the output, and a route of correction. The cases were selected because accessible primary materials expose a contrast relevant to one of these relationships. Unequal national coverage is deliberate. The selection favours documented disputes and published rules; it cannot establish the prevalence of unreported use, compliance or harm.

Comparisons completed in this book

The following register states the role of each focal jurisdiction. It records legal comparisons, not a count of evaluated deployments. Bangladesh appears only as the setting of Hoque's contextual scholarship; European procurement clauses and international standards serve as reference instruments rather than additional national case studies.

Jurisdiction and principal materials	Comparative role	Evidence boundary
South Africa: Shilubana, Constitution, PAJA, POPIA and evidence legislation	Living customary law; administrative correction; data rights and responsibility	Judicial and statutory analysis; constructed benefit-file example, not an observed deployment
Kenya: Constitution arts. 159, 165 and 170; Data Protection Act	Conditional religious jurisdiction; referral and institutional authority; data-protection oversight and impact assessment	Constitutional and documentary analysis; surname-based referral is constructed
Colombia: T-323/24 and Agreement PCSJA24-12243	Judicial reasons, disclosure and the move from a constitutional judgment to implementation rules	Specific judgment and subsequent rule; no claim of uniform court compliance
Brazil: CNJ Resolution 615/2025, LGPD, CDC and ANPD Resolution 19/2024	Contrast in judicial governance, data protection, transfer mechanisms and private-law responsibility	Instrument-specific analysis; later changes identified where checked; constructed private-service dispute
Pakistan: BISP Act and programme evidence; Sarki judgment	Allocation of programme authority and correction; treatment of video evidence and proof	Historical programme documents and one judgment; not a present-day audit of all BISP decisions
India: electronic-evidence legislation and commencement	Electronic proof and transition between statutory regimes	Doctrinal comparison; no inference of measured accuracy or nationwide uptake

This selection spans distinct institutional questions and supplies contrasts within them. Colombia and Brazil are paired on judicial governance, South Africa and Brazil on data-related correction and liability, and India and Pakistan on the legal treatment

of electronic evidence. Kenya and South Africa are paired to show why legal pluralism produces different source and jurisdiction problems. A country's appearance in one comparison does not establish equivalent coverage of its entire legal system.

The country choices also have costs. Publicly available material and English-accessible scholarship make some institutions easier to examine. Portuguese and Spanish official instruments were consulted for the propositions identified in the text; English explanations of their meaning are the book's analysis unless an official translation is specified. South African, Kenyan, Indian and Pakistani materials were consulted in available English versions. The study does not claim independent verification across all indigenous, regional or minority-language legal sources. This limits its ability to describe lived legal practice and reinforces the importance of preserving dispute and uncertainty in digital records.

Source discovery and verification

Sources were checked as of 20 October 2026. Discovery was directed by the book's claims and comparative questions: named judgments and statutes, institutional author and instrument number, article title and DOI, and combinations of the relevant legal function with the jurisdiction. Searches followed official court, government and regulator repositories, publisher records, DOI-linked metadata and author or university repositories. This was a targeted search and reference audit, not a preregistered systematic review; no search-result counts or exhaustive database coverage are claimed.

Inclusion required a source identifiable by author or issuing body, title or legal identifier, date or version, and a retrievable locator. Primary legal material was preferred for legal propositions. Publisher abstracts could establish a work's identity and stated topic but were not used as substitutes for inaccessible doctrinal analysis. Selected full-text passages were checked for the propositions they support. Older preprints were retained where the argument relied on that examined version and a final publication had not been compared in full. The references identify the version actually used.

Internal compilations and drafting notes were excluded as substantive authorities. Duplicated reference entries were merged by underlying work, preserving legally material versions. An unavailable page was treated as an access limitation, not proof that the source was fictitious. Claims dependent on unpreserved interface observations or unchecked detailed content were removed, narrowed or replaced by identifiable primary material. These decisions distinguish bibliographic authenticity from evidential adequacy: a real source can still be the wrong support for a proposition.

Legal version checks were tied to the proposition and date under discussion. The analysis distinguishes enactment, commencement, amendment and historical application, and uses official consolidated texts where available. Dated original instruments remain appropriate for historical questions. Later rules are considered where they ma-

terially change an earlier account, as in Colombia's judicial AI implementation and Brazil's international-transfer regulation. The research date is a source-checking boundary, not a claim that every law or institutional practice in the six countries has been exhaustively updated.

Kinds of claim and limits of inference

Four kinds of proposition are kept separate. Legal propositions concern the requirements and reasoning of identified authorities. Documentary observations concern what a published institutional record states. Empirical findings are attributed to the study that produced them, with its population and period. Normative proposals and constructed examples develop the book's argument. A specification cannot prove delivery, a policy cannot prove compliance, and a constructed scenario cannot prove that a named institution actually operated the described system.

The original legal analyses follow a common sequence: identify the relevant act and interest; state the controlling provisions or judicial reasoning; apply them to the facts available; address a plausible counterargument; and explain the result and its limits. Constructed examples hold the technical task constant while varying its procedural consequence, or hold the error constant while comparing remedies. Their purpose is analytic clarification. All numerical workloads, costs and performance assumptions used in those examples are illustrative unless expressly attributed to an empirical source.

A source was assessed at the level of the claim. A judgment was checked for the question decided, material reasoning and disposition; a statute for its operative provision and relevant temporal status; a scholarly work for the passage actually used. A supporting citation establishes neither the truth of every statement in that source nor the generalisability of its setting. Pinpoints in the text and the consolidated references allow readers to test the material inferences.

The principal selection limitation is visibility. Published litigation and institutional rules make contested or formalised uses easier to discuss than routine, informal or undisclosed practice. The book consequently establishes legal conditions and plausible mechanisms, not frequencies. It does not infer project failure from a missing update or successful adoption from an announcement. Claims about user experience, compliance and distributional effects require operational evidence beyond this desk study.

Future empirical testing

The argument generates testable questions. Does an automated label alter which exceptions an authorised officer considers? Can a corrected source record change a downstream decision in time to matter? Does a formally optional output become compulsory under workload pressure? Future studies should compare the same legal function, population and period, documenting concurrent staffing, records and procedural changes. Before-and-after timing alone would not identify AI's effect.

Fieldwork would require its own sampling, ethical review, consent, confidentiality arrangements and language competence. It should include affected people who could not complete the digital route, frontline staff and officials able to change outcomes, rather than only project sponsors. Disagreement between their accounts should be investigated through records and institutional incentives. These are proposed extensions, not activities claimed as completed for this book.

The six jurisdictions should not become a compulsory quota for later work. A smaller comparison with accessible case records, lawful data access and a credible account of institutional practice can answer more than a nominally balanced collection of promotional descriptions. Non-adoption, suspension and redesign are relevant where their reasons are documented. Missing evidence should remain a stated limit on a particular inference.

This method supports a conditional comparative conclusion: the same computational function can acquire materially different legal significance through its institutional consequence, source structure and corrective route. It does not support a ranking of countries, a pooled estimate of effectiveness or a claim that one regulatory model should be transplanted intact. The instruments that follow make the conditions of an application explicit so that the comparative argument can be challenged and refined.

Table 1.4 is the reusable extraction instrument underlying this method. The case description is maintained once, while each consequential legal or empirical proposition has its own source and support record. It can also structure a later empirical study, provided that the study's new sources, dates and limits are recorded separately.

Table 1.4. Comparative case-selection and evidence-extraction template

Field	Information to record	Check before drawing a conclusion
Case identity	Country, institution, territorial setting, legal function, affected population, and case identifier.	Is the institutional unit defined rather than inferred from a national label?
Selection rationale	Research question, functional family, institutional contrast, alternatives considered, and inclusion reason.	What meaningful difference can this case help explain?

Table 1.4. Comparative case-selection and evidence-extraction template

Field	Information to record	Check before drawing a conclusion
System description	Task, inputs, outputs, provider, model and configuration, users, and connected processes.	Are AI, ordinary automation, and the institutional decision role distinguished?
Deployment status	Proposal, demonstration, pilot, routine use, suspension, discontinuation, or replacement; dated source.	Does evidence establish status, or only an announcement?
Authority and legal status	Instrument, issuing body, exact provision, scope, relevant interpretation, and status.	Binding rule, proposal, guidance, contract, technical standard, or unresolved?
Chronology	Enactment, publication, commencement, amendment, system change, and observation dates.	Which legal and system versions applied to the event studied?
Language and provenance	Original language, authoritative text, translation source, access date, and interpretation limits.	Could an omitted qualification or translation uncertainty change the conclusion?
Workflow and discretion	Entry, classification, routing, review, decision, notice, correction, and powers to depart.	What demonstrates actual practice rather than intended procedure?
Records and coverage	Dataset origin, category definitions, included and excluded practices, populations, and access restrictions.	What is missing, and how does that limit the claim?
Baseline and comparator	Prior process, staffing, records, case mix, concurrent reforms, and feasible non-AI alternative.	Is the comparison documented and appropriate to the legal objective?
Outcome evidence	Indicator, period, denominator, population, distribution, uncertainty, and assessment method.	Does evidence support the scale of the claimed change?
Claim-level source record	Exact proposition, source, page or paragraph, support type, qualifications, and counterevidence.	Requirement, practice, evaluation, promotion, or authorial inference?

Table 1.4. Comparative case-selection and evidence-extraction template

Field	Information to record	Check before drawing a conclusion
Ethics and limitations	Access conditions, consent where relevant, confidentiality, missing material, and disagreements.	Can the material be used, and must the conclusion be narrowed?
Review and update	Reviewer, outstanding questions, observation cutoff, version changes, and next review event.	What new evidence would alter the case account or its comparative use?

The template supports exclusion and qualification as well as inclusion. A case can remain useful for understanding authority while offering insufficient evidence of performance. Another may document an operational change while leaving the full legal basis unresolved. The final comparison should specify that difference. Its credibility depends on preserving the boundaries of each claim, not on making every case appear equally complete or forcing all six jurisdictions into a single explanatory narrative.

1.5 Criteria for Evaluating Legal and Institutional Change

A project's success cannot be established solely by installation, staff adoption, processing volume, or a reduction in the time needed to produce an output. A legal institution must also consider whether the output serves an authorised function, whether the person affected can participate, and whether errors can be corrected before their consequences become irreversible. Evaluation therefore begins with the problem and the existing process, not with a preferred model. The institution needs to know what would count as an improvement before a demonstration makes one solution appear inevitable.

The framework developed here separates legality, accuracy, procedural fairness, accessibility, cost, institutional control, and correction. NIST's AI Risk Management Framework supplies a useful technical point of reference through its voluntary, context-sensitive approach and its governance, mapping, measurement, and management functions (NIST, 2023, p. 2; Part 2). It is not domestic law in the focal jurisdictions. The legal and institutional criteria below are the chapter's analytical framework, to be completed against the applicable duties and conditions of a defined use rather than converted into a universal numerical score.

The baseline as an account of an existing legal process

A baseline describes the process as it operates, including both difficulties and safeguards. It identifies entry routes, required records, decision criteria, responsible offices, notice, opportunities to explain, and available challenge. It also records where people fail to enter, abandon an attempt, or need assistance. Counting completed applications alone cannot describe the experience of people whose attempts never become cases. The baseline population and the proposed system's intended population must therefore be distinguished before their outcomes are compared.

Time requires more than one measure. Institutional processing time is different from the person's time spent obtaining documents, travelling, waiting, seeking advice, or correcting an error. A shorter staff interaction may shift work to the applicant rather than remove it. Equally, an apparently slow step may provide an opportunity to clarify a disputed fact. Evaluation should distinguish avoidable delay from time performing a necessary procedural function. The aim is not to preserve every existing step, but to identify what would be lost as well as what would be saved by removing it.

The unit of observation must be specified. A document processed, an application completed, a decision issued, and a remedy obtained are different events. An institution can increase the first without improving the last. Outcome measures should follow the relevant legal objective and retain a connection between intermediate activity and final consequence. A claim about faster resolution should state where the clock begins, where it ends, which matters are excluded, and whether later correction or reconsideration falls outside the reported period.

Baseline uncertainty needs its own response. Missing data should not be replaced with favourable estimates chosen to make a future improvement measurable. The institution may need an initial records exercise, a narrower pilot, or a limited claim about the outcomes it can observe. A clearer account of the existing process may reveal that the primary problem is inconsistent recording, confusing forms, or an unstaffed correction route. In that situation, the baseline is not merely preparation for AI; it can change the choice of intervention.

Legality and the limits of balancing

The legal-authority assessment in Section 1.3 is a distinct part of evaluation. It identifies the powers and conditions applicable to procurement, data activity, procedure, and decision-making. An anticipated efficiency gain does not establish a missing authority. Within this framework, an unresolved essential permission cannot be counted as a positive finding because the expected benefits are large. The exact legal consequence of unauthorised conduct remains jurisdiction-specific; the analytical point is that predicted utility and demonstrated competence answer different questions.

Legal assessment also concerns the criteria operationalised by the system. A model may reliably predict a characteristic that the institution cannot lawfully make decisive.

A procedure may collect information that improves technical performance but is not justified for the function concerned. These are not necessarily errors in computation. They are questions about the relationship between the system's objective and the institution's permissible purpose. The evaluation record should identify that relationship explicitly rather than treat accuracy as a substitute for legal relevance.

Some choices remain legitimately open after minimum legal conditions have been established. An institution may compare different lawful designs, levels of assistance, and allocations of resources. The evaluation should explain those choices and their distributional consequences. It should not imply that every criterion is interchangeable, or that a high average benefit cancels the removal of a required safeguard. A reasoned account of trade-offs is more informative than a composite result that conceals which conditions were satisfied and which remained unresolved.

Function-specific accuracy and reliability

Accuracy requires a defined task and reference standard. Search evaluation examines relevant material retrieved and omitted; classification evaluation examines false inclusions, false exclusions and their consequences. Prediction requires a defined target and an account of its relevance to the task. Translation must preserve legally material meaning. Generative outputs require checks for unsupported claims, omissions and fidelity to sources. A single aggregate percentage cannot, by itself, convey these different kinds of error or their legal consequences.

False positive and false negative errors have differing implications. For example, in a theoretical legal aid intake scenario, misidentifying a complete application as being incomplete could result in either delays or denial of service. Mistakenly treating an absent required document as attached could leave an incomplete file undetected until a later stage. The significance of each is contingent upon the operational workflow; the opportunity to correct the error; and the impact of the error on the individual involved. Thus, testing should provide documentation of events that occur after an error is made (not simply whether the output matches a label). A preliminary warning and an automatic closure can have different consequences even when both originate from the same classifier.

A review of the reference standard is warranted. A prior decision may include errors or may reflect aspects of a case not included within the training set. There may be reasonable disagreements between reviewers relative to interpretations. Therefore, a review must differentiate clearly between material that is objectively incorrect (e.g., non-existent citations), and disputed legal conclusions. Furthermore, a review must describe the sources and level of expertise that were relied upon to make determinations regarding correctness; allow for the documentation of reasonable disagreements; and prevent an assumption that an agreement exists with one annotator constitutes a total validation of legality. If the issues cannot be resolved through further means, then the permissible uses for that tool may need to be similarly restricted.

In addition to evaluating the outputs of a model, reliability encompasses the environment in which a valid output can be generated and reviewed. Issues such as missing records; unavailability of services; changing input values; or inability to access the original may preclude production of valid outputs using an otherwise capable tool. The bounds of the relevant claim are dependent upon the version of the system utilized; the available data; institutional policies and procedures; user interaction(s); and time frame observed. Success under specified conditions would not demonstrate reliability for all offices that receive the output. The evaluation should include statements describing the conditions tested; and provide a pathway for consideration of events that occurred outside those described conditions.

Procedural fairness and effective review

Procedural fairness is examined through the person's opportunity to understand and respond, and the reviewer's ability to reconsider. The applicable law determines the precise requirements. The evaluation asks whether the proposed arrangement actually supports them. A notice may be formally delivered but fail to explain what needs correction. A complaint channel may exist but provide no access to the evidence used. A reviewer may be named but lack authority to change the result. The presence of each component is different from its effective operation.

Meaningful review can be tested through a reconstructed case. Can the reviewer identify the input used, inspect the original material, determine which rule or configuration mattered, and understand the action taken? Can the affected person present relevant information that did not fit the original fields? Can reconsideration change the operative outcome? These questions connect technical records to institutional powers. An explanation generated after a complaint is insufficient if it cannot be related to the actual process that produced the disputed result.

The timing of review matters. A correction obtained after a deadline or a lost opportunity may not restore the previous position. The institution should identify situations in which reliance can be paused, further information accepted, or an alternative route used while the dispute is examined. These are proposed design questions whose legal implementation requires the appropriate authority. They prevent evaluation from assuming that a later right of challenge makes every earlier burden harmless.

Accessibility and the distribution of benefits and burdens

To assess both organisational (institutional) and user (affected-person) results you need to record these individually. While reduced costs will certainly assist the organisation; the increased administrative effort required of users is a potential drawback. On the other hand, an increase in staff time could provide users with greater opportunities to correct errors or avoid multiple journeys. However, no assessment can be made from either the cost per licence nor solely from the volume of licences.

Your relevant account would need to determine which party receives the benefit; which party bears the burden; and if your sample group contains individuals who have difficulty using the main electronic route.

Assessing the completeness of the legal process is not simply a matter of whether a user can open the interface. To complete the legal process, the user must be able to read / understand the requirements of the legal process; communicate all relevant information; obtain assistance when needed; receive a meaningful response to their actions; and be able to pursue correction. If this is lawful and provided there is sufficient supporting evidence, assessment should evaluate any differences across language; location; gender; disability; and documentation status. Group definitions should be justified, and the analysis should not assume that everyone within a group has the same experience. Similarly, you should not allow your desire to measure inclusivity justify unnecessary collection of sensitive information.

As previously mentioned, Alternative Channels must be evaluated practically. Although you may consider a nominal offline channel to be available, it may still depend on travel; longer queues; and/or an unavailable official. In addition, an Assisted Digital Channel can be successful if support is readily available and maintains the individual's control over their account. While functional equivalence does not mean equivalent processes, functional equivalence means providing an honest description of any burdens placed upon users as well as providing similar opportunities. A system that appears to operate efficiently due to diversion of difficult cases without providing equal service has not shown the same level of improvement as a system that assists difficult cases in reaching a lawful result.

Although Consultation can identify issues with the design that were not identified by the Project Team, Consultation cannot necessarily validate the design. Individuals participating in consultation may differ regarding what they perceive as convenient, private, assistive, or acceptable levels of error. The views of participants should be documented along with the response of the institution. User Satisfaction is good evidence of how users experienced their interaction with a particular system; however, user satisfaction is not evidence of legitimate authority or reliable results. Users who do not interact with a system at all deserve consideration since their absence from Satisfaction Data cannot be considered reasonable acceptance.

Costs: Capacity and Realistic Non-AI Alternatives

Total Process Costs include development of records; integration; employee training; operation; validation; accessibility support; monitoring; correction; Incident Handling and future transition. Total Process Costs should detail known expenditures versus estimated unknowns and indicate which parties incur such costs. As noted above, a low-cost pilot may rely on resources that are not available when the Service is expanded. Time saved during drafting may be offset by the time spent reviewing and correcting. The comparative analysis should compare total expenditures incurred over a given timeframe and include expenditures related to sustaining a legally compliant service (i.e., operating software), not just keeping software operational.

Capacity refers to demonstrating the ability to accomplish those tasks assigned. An appropriate Safeguard will contain elements such as: competent personnel; time to devote to the Safeguard; documentation to support the existence of the Safeguard; and authority to act under the terms of the Safeguard. Independent checks are often critical components of the design. Therefore, independent checks must be funded, not assumed. The issue is not whether the organisation has established policies that contain Safeguards, but rather whether the implementation of the proposal provides sufficient feasibility for implementing those policies. Shared Services or External Expertise may provide some assistance; nevertheless, allocation of responsibility for evidence production must continue to be clear.

An alternative solution does not consist exclusively of "do nothing". In fact, employing additional employees; improving indexes; creating clearer forms; developing reliable source documents; simplifying routing rules; and enhancing communication may be viable solutions for addressing the same issue(s). All possible arrangements should be assessed against comparable objectives for compliance with law and realistic limitations. A non-AI solution is not inherently superior; nor should the existing process be assumed acceptable. The comparative study will seek to determine which achievable arrangement modifies accuracy, access, costs, institutional control, and correction.

A legitimate concern is that evaluation may prolong serious deficiencies in current practices. This concern is legitimate when delay itself poses risk. Urgency can justify limited, timely comparisons and explicit decisions under uncertainty; it cannot establish the reliability of a preferred technology. Urgent needs for change may permit expedited assessments, rapid trials or immediately implementable non-AI enhancements. Urgency neither guarantees that untested consequential applications will generate stated benefits nor that resultant errors will be more easily correctible than those currently extant.

Correction, continuation, and the decision not to proceed

Correction is an outcome, not simply a policy statement. The institution needs to distinguish amending a source record, preventing further reliance on it, communicating the change to recipients, reconsidering the affected decision, and providing available relief. Repairing a classifier does not necessarily restore a wrongly closed application. A corrected source may coexist with an unchanged downstream result. Evaluation must follow the relevant chain far enough to establish whether the person's problem has actually been addressed.

A pilot needs a defined scope, observation period, comparator, accountable approval and stopping conditions. Its outputs and consequential changes must be reconstructable. Criteria should be set before results are known and should reflect the task's legal consequence, reversibility and capacity for correction. The worked example below shows how uncertainty can determine a restricted decision without treating an invented performance number as a legal standard.

Evidence may support continuation, conditional continuation, restricted use, redesign, further inquiry, or rejection. The reasons should identify the problem rather than attach a general label to the technology. Relevant findings include conflict with applicable law, an unresolved essential authority, consequential errors that cannot be controlled, an inaccessible remedy, inability to operate necessary safeguards, or a feasible alternative that better addresses the identified need. Some deficiencies can be remedied; others defeat the proposed use in its present form. The record should distinguish those conclusions and the conditions for reconsideration.

A non-aggregated evaluation matrix

Table 1.5 records each dimension separately. It does not allocate universal weights or produce a national or institutional readiness score. The matrix is completed from baseline and assessment evidence, retaining limitations and disagreement. Its purpose is to explain the relationship between the proposed application, the institution's legal duties, and the outcomes experienced by affected people.

A worked choice between three intake designs

Assume a small advice office must process 200 applications a week and has 20 staff hours available for intake and correction. These are constructed planning figures. The governing procedure, also stipulated for the example, requires consideration of explanations for unavailable documents and permits no automatic closure solely on a missing-attachment flag. The office considers manual processing with clearer forms, AI flags with full human review, and automated closure after a flag. The legal condition is applied first: the third option is excluded regardless of its apparent speed.

For the two eligible designs, use explicit workload assumptions. Clearer forms require five minutes of staff time per application and two hours a week for correction: 1,000

minutes plus 120 minutes, or 18 hours 40 minutes. AI flags with review require four minutes per application, three hours for checking exceptions and complaints, and two hours for maintaining sources and monitoring: 800 minutes plus 300 minutes, or 18 hours 20 minutes. On these assumptions the AI design saves only 20 minutes weekly. Acquisition cost has not yet been included, and neither calculation establishes actual service quality.

A sensitivity check changes the decision. If exception checking takes five hours rather than three, the AI design requires 20 hours 20 minutes and exceeds available capacity. If clearer forms reduce processing to four and a half minutes, the manual design requires 17 hours and is faster. These are not predictions; they show which uncertain quantities matter. The next useful evidence is observed total handling time, reasons for exceptions, missed cases and correction time under a trial that imposes no adverse automated consequence. A vendor's drafting-speed claim cannot resolve the comparison.

The decision on the stated evidence is to adopt clearer forms and retain manual intake while testing AI suggestions in a limited setting, if the office can fund that test separately. It is not to divert required casework capacity to an unproved saving. A later decision could change if the trial demonstrates a meaningful net benefit without reducing access, considering the costs of maintenance and correction. No universal accuracy percentage is needed to reach this provisional choice: legality eliminates one design, capacity constrains another, and sensitivity analysis identifies evidence capable of changing the conclusion.

Distribution remains a separate test. Even a large average time saving would not justify a system whose errors leave a language group unable to explain missing records. The office should assess performance against the relevant procedural opportunity, using enough evidence to support the proposed scope. Where that evidence is absent, restriction to an internal aid may be justified. This sequential method preserves non-negotiable duties while still allowing reasoned choices among lawful alternatives.

Table 1.5. Evaluation matrix for a proposed legal AI application

Dimension	Baseline evidence	Assessment or pilot evidence	Reason to reconsider the proposal
Legality	Mandate, assigned office, procedure, data permissions, and applicable review rules.	Completed activity-specific authority analysis; verified conditions and scope.	Conflict with applicable law or unresolved authority essential to the use.

Table 1.5. Evaluation matrix for a proposed legal AI application

Dimension	Baseline evidence	Assessment or pilot evidence	Reason to reconsider the proposal
Accuracy and reliability	Existing errors, source quality, case mix, consequences, and reference standard.	Task-specific results; omissions; uncertainty; expert disagreement; tested operating conditions.	Consequential errors uncontrolled, critical performance untested, or legal relevance unestablished.
Procedural fairness	Notice, information, response, reasons, exceptions, and reconsideration in existing practice.	Person and reviewer can understand, challenge, reconstruct, and change the relevant result.	Required participation or effective review is removed, obstructed, or not demonstrated.
Accessibility and distribution	Entry failures, abandonment, assistance, alternative channels, user burdens, and relevant differences.	Ability to complete the task and correct errors across tested languages, channels, and circumstances.	Apparent gains depend on excluding difficult cases or leaving affected groups without a workable route.
Cost and capacity	Staff and user costs; records work; existing competence; realistic alternative arrangements.	Whole-process expenditure, review time, assistance, monitoring, correction, and transition needs.	Necessary safeguards cannot be operated or resources cannot sustain the approved use.
Correction and remedy	Detection, source correction, propagation, reconsideration, and available relief.	Time and evidence needed for correction; responsible actors; downstream repair; unresolved consequences.	Fixing the output leaves the affected legal position uncorrected or review is too late to be effective.
Control and continuity	Responsibility, record access, service dependencies, and existing fallback.	Approval of changes; usable records; power to pause; tested fallback and exit arrangements.	Institution cannot supervise, investigate, stop, or replace the system while maintaining its function.
Alternatives and public outcome	Defined legal objective; feasible staffing, records, procedural, and non-AI options.	Comparison using the same population, period, costs, safeguards, and outcome measures.	No demonstrated advantage for the proposed use, or another feasible design addresses the need better.

The competent institutional authority should record the decision, supporting evidence, permitted use, monitoring responsibilities, and review triggers. Technical updates, altered data, a new population, a different legal criterion, or changes in human

control may require reassessment. Evaluation thus continues beyond acquisition. The enduring question is whether the institution can maintain lawful, intelligible, accessible, and correctable practice under the conditions in which the system is actually used.

Chapter 1 Key Arguments

The first concern relates to the nature of change. While the adoption of AI is not sufficient evidence of legal change, and legal change does not always signify an improvement, legal change is understood as a material change in authority, interpretation, discretion, professional responsibility, participation, or access to remedies. There are many examples of technologies improving existing processes of administration without legally changing anything. Similarly there are many technologies that will effectively change the way the law operates in practice with no corresponding formal change in the way the law allocates power. An appropriately robust understanding of legal change must identify both the mechanisms by which legal change occurs and the consequences that arise from those changes. In this regard, determining whether an application has produced legal change based on the complexity of the technology alone would be inappropriate.

Secondly, legal treatment should reflect the functional characteristics of the institution and its position within that institution. Digitalization, searching, categorizing, predicting, translating, and generating different functions contribute differently to tasks such as identifying parties to litigation, and each produces different types of errors. Therefore, combining these functions in one interface does not eliminate their differing functional attributes. A prediction does not become a legal standard simply because it is accurate; a generated explanation of a reason for doing something is not evidence of the actual reason that was used; and a categorization that determines who is eligible for consideration in a proceeding is not merely administrative.

Assistance becomes meaningful only when the individual responsible for making decisions using the technology can review the materials upon which the recommendations were made; exercise judgment regarding those recommendations; elect to deviate from them; and have that deviation recognized as valid.

Thirdly, systems' records represent institutionalized perceptions of reality rather than complete accounts of all aspects of law. Legal recognition, social authority, practical accessibility, and digital visibility must be differentiated. Therefore, the comparative study must examine which legal practices and descriptions are recorded; which are reframed; and which continue to exist outside of any record; and why. Historical explanations provide context when they demonstrate a relationship between past concepts and current categories or authority structures. However, historical explanation cannot serve as a substitute for contemporary domestic analysis. Nor should either formal institutions or informal customs, religion, or community-based arrangements be roman-

ticized; and that additional data is not automatically the best solution for unmet knowledge.

Fourthly, authorization is specific to both an action and its resultant effects. The authority granted to purchase software does not imply authority to reveal confidential information; derive an attribute; modify procedures; or determine an outcome. The relevant documents and individuals having lawful authority over each step of a process must be identified. Analysis of delegations must consider the purpose that is assigned to a delegatee by law, as well as the amount of control that is exercised over delegated activities in practice. Authority may already exist for certain uses under existing laws; however other uses may require a modification to an existing legal instrument or creation of a new legal instrument. Finding missing authority, ambiguous authority, unauthorized actions taken in reliance upon an authorized actor's directions, and insufficient capacity are four distinct conclusions.

Lastly, the comparative studies conducted at completion are bounded legal analyses rather than six separate assessments of national deployments of technologies. Each study distinguishes among legal requirements, documentation provided in support of statements made about technology use; factual research results; and suggestions contained in this publication. The rationale for selecting particular jurisdictions to address a specific issue acknowledges varying levels of document availability for purposes of comparative analysis. Thus, limiting regional and causal claims allows for providing a rational assessment of how authority, representation, and redress affect the legal implications of performing a computational task.

Evaluating legal impacts of technological innovations requires linking improvements in institutional performance to positive outcomes experienced by people impacted by technology use. Entry failure rates; user burden; and operational corrections are important baselines for measuring success. Measuring accuracy requires defining accuracy in terms of the task being performed and its related legal effect(s). The existence of resources and authority to implement safeguards; and costs associated with implementing safeguards including verifying outputs; assisting users; processing complaints; monitoring usage; and allowing users to exit are also relevant factors. Non-AI alternatives to AI solutions should be included in comparable evaluations. Providing a thoughtful decision-making framework supports evaluating the potential adoption / implementation of AI solutions; restricting their use; modifying their design; delaying implementation until further information is available; or rejecting proposed implementations entirely without reducing evidence to a single numerical score.

CHAPTER

02

AI in Legal Practice and the Administration of Justice

IN THIS CHAPTER

2.1 Legal Research, Drafting, and Document Review

2.2 Court Administration and Case Management

2.3 Judicial Reasoning and AI-Assisted Decision-Making

2.4 Legal Aid, Language Access, and Public Legal Information

2.5 Professional Duties and Changes in Legal Work

Introduction

Artificial Intelligence (AI) enters legal practice via tasks that have some degree of accountability. A research assistant identifies authorities; a lawyer translates an instruction into a pleading; a Registry assigns a category to a filing; an interpreter transmits a party's version of events; and a Judge provides reasoning as to why one argument prevails and another fails. While software could potentially alter the rate and manner in which these activities occur, such modifications would not necessarily confer upon the software the authorization or professional obligations associated with those activities. Therefore, the central issue is not how much work a given system can generate, but rather whether that work is acceptable to the appropriate parties and institutions, i.e., verifiable, accepted, communicated, contested and corrected by persons or entities authorized to take action.

Chapter 1 described the distinction between technology adoption and legal transformation. This chapter extends that distinction to both the generation of legal work and the administration of justice. For example, a tool may remain simply an aid to an existing task and thus continue to perform that task at varying speeds and with varied distributions of effort, however, when the output of that tool determines whether a document proceeds to a reviewer or whether a particular individual must take further action, then that tool becomes significant. Thus, a seemingly minor categorization may result in an urgent filing being delayed; a well-articulated summary may fail to include an essential factual element necessary to support an exception; and a rationale may be articulate in supporting a position, but incorrectly attribute that position to an authority that has never made that determination. What gives rise to the legal significance of AI-generated output is not its characterization as "an assistant," but rather the relationship between the output and the resultant act.

This chapter's legal foundation consists of local laws regulating actors whose work is addressed herein. Two Colombian sources were selected - Constitutional Court judgment T-323/24 and Judicial Council Agreement PCSJA24-12243 - for comparative purposes with Brazil's CNJ Resolution No. 615/2025 ("Resolution"). In addition, the professional codes of South Africa's Legal Practice Council and Brazil's OAB provide additional comparative material regarding lawyers' independence, confidentiality and responsibility. Because this chapter tests universal safeguard provisions against specific responsibilities rather than deriving domestic approval based upon global aspiration, these sources provide a means to examine legal design. It is also important to note that compliance with an instrument does not demonstrate that all institutions comply with that instrument.

Although each country in the book's focus group - Pakistan, India, Kenya, South Africa, Brazil and Colombia - represent distinct legal environments, this chapter conducts comparative analysis using similar comparisons across countries rather than conducting an exhaustive list. The primary judicial comparison held preparation of a

summary of the record constant while comparing permissions granted by the jurisdiction, human involvement in controlling the system, disclosure requirements and record retention requirements. The professional comparison was conducted similarly, holding confidential drafting constant. Identified provisions - not assumptions about characteristics of the Global South - explain differences between jurisdictions.

This chapter presents five interrelated arguments. First, verified drafting requires identification of a direct link between a proposition and its source as well as identification of the case in which it will be employed. Second, court administration requires correction processes contemporaneously with classification, allocation or scheduling decisions affecting access. Third, judicial assistance requires control over the evidentiary and legal bases of judgments as opposed to only signing off on generated rationales. Fourth, public legal services require pathways connecting information to competent assistance as well as remedies available to consumers. Fifth, professional regulation requires identifiable names of individuals with demonstrated competence as well as credible allocations of time for supervisory activities.

These arguments lead to a practical legal-practice/court-use protocol. Five related tables describe the work subject to verification / review, the person reviewing / verifying it, the data maintained for evidence of prior versions, and the conditions under which escalating corrective actions are warranted. As noted above, each table should be considered in conjunction with others. For example, even though a reviewed draft may pass muster as drafted, if the wrong version is filed for consideration - e.g., prior version vs. current version - the draft fails. Similarly, even though a registry record is correctly updated, if a hearing notice is not re-notified due to incorrect or incomplete data, the record fails. Finally, even though a referral may meet formal criteria for completion - e.g., proper documentation submitted - failure to obtain acceptance from a provider results in referral failure. Accountability relates to maintaining continuity throughout the entire process including hand-offs between teams / institutions.

Each of the tables constitutes an instrumental guide created by authors for application of the methodology presented above. Each completed protocol should be linked to the shared activity, matter, source-record and output-version identifiers, together with the responsible actor and review date. Fields are only included where relevant to the task. Table 2.1 associates propositions with their respective verified-sources; Table 2.2 follows an administrative procedure as well as notification; Table 2.3 describes permissible-judicial functions; Table 2.4 traces the user's next-step; and Table 2.5 allocates professional work. Chapter 3 will add fields describing the processing-purpose, recipient and correction-event.

2.1 Legal Research, Drafting, and Document Review

Verification concerns propositions, not the appearance of a document

A legal proposition is adequately verified once there is sufficient evidence available to allow an accountable reviewer to identify the reasons why the proposition was supported (why the proposition's support exists), and how it applies or relates to the specific case being worked on (the impact of that support). Citing a plausible source, using an accurately quoted passage, and developing a good structure to present this information are each answering somewhat different types of questions. The question about where you got your information addresses the identity of the source you are citing. The question about what that source said answers what the source stated. The question of how relevant and effective is that statement within a particular legal context is addressed by the argument. Therefore, if we allow success at one level to satisfy requirements at all three, then verification will be inadequate.

There are two aspects of verification: the epistemological aspect and the professional / technical aspect. The epistemological concern is whether a proposition is supported. The professional concern is whether the person who is accountable for verifying a proposition has conducted an inquiry commensurate with their intended use of the verified proposition. If a lawyer arrives at a correct answer through accident and not because he or she used a reliable process to arrive at his or her conclusion, it does not necessarily mean that they have developed a consistent and reliable process. On the other hand, an incomplete answer may serve as a starting point for further research by providing leads that can be investigated in greater detail. It is essential, however, that any such incompleteness remain evident so that a lawyer may complete the necessary research before relying upon those conclusions. In order to evaluate an attorney's suggestions appropriately, a distinction must be drawn between the quality of any preliminary ideas or suggestions an attorney provides as compared to the quality of the attorney's completed product.

Therefore, there should be three separate stages or levels of review to ensure adequate verification. First, source verification ensures that the materials relied upon to establish support exist. Second, legal verification establishes the meaning, extent and application of that support. Finally, document-level review verifies whether individual propositions established separately based upon independent sources were properly compiled and presented as a coherent account of applicable law. For example, a pleading may include references to authoritative decisions but fail to include some critical limitation that renders them irrelevant to the primary issue in dispute. Similarly, a contract may include standard provisions but assign risks to clients which were never accepted by those clients.

Existence, currency, jurisdiction, and the actual holding

To begin the existence check process you need to retrieve and read the underlying authority from the issuing institution, an official repository, the publisher or a reliable law report. Take note of the title, who issued it, when it was issued, what identifiers are associated with it and take some notes about the passage. You may use search snippets, summary generated articles or secondary reference materials to help find the information. However, do not consider these sources to replace the original documents upon which your claim (consequential) is based. If you are unable to obtain access to an authority document do not indicate you have personally reviewed it. You can either remove the unresolved proposition from the draft, qualify the proposition or refer it back to someone else for additional research prior to releasing.

For currency purposes you will need to look at the time frame applicable to the law question being asked. The latest version of a statute does not necessarily make it the best applicable statute. There could be an older transaction / previous proceeding that incorporates a previous version of a statute or a transitional provision. Similarly there could be an older case where the language used in the case has since been altered. Your review should include identifying all changes / amendments to statutes including any commencement provisions and any later treatment of previously treated propositions. Also, you will want to verify if any temporal limitations were placed on the argument. While the system's statement indicating it has current access to a resource is not considered an entry into this verification process; you will still need to verify what resources were actually accessed and when they were accessed.

In addition to just matching a country name jurisdictional relevance also includes factors such as the type of court or tribunal hearing the matter, geographic area within which the matter is heard, subject matter of the dispute / proceedings, type of procedure being followed, and the authority the relevant court must or may give to the cited source. Comparative decisions can provide guidance but may not be binding. A procedural decision made by a specialized tribunal may not apply to an ordinary civil proceeding. In each instance, counsel must describe both the status in which a source is being used and the basis for comparison. Where that status is unclear, the ambiguity must be discussed in the analysis as opposed to simply assuming confidence in one's characterization.

A difference exists between the actual holding of a decision and a party's submissions regarding those issues, quotations from other authors / sources, illustrative comments or results that are limited to only certain facts. To distinguish between these alternatives, it is essential that you "read around" the quote/passage under discussion. You will want to determine: What issue(s) were actually decided? What rationale did the decision-maker require to reach their conclusion? Were there any limits or qualifications accompanying that decision? When addressing legislative matters, you will want to perform similar tasks by reading relevant definitions, exclusions / exceptions contained within the statute, related provisions / statutes and any applicable commence-

ment conditions. Isolated sentences disconnected from their surrounding framework can be technically correct while being legally inaccurate.

By way of example, suppose a generated draft cited a genuine judicial opinion as support for the proposition that an application is always admissible. As noted above however, in reality the trial judge allowed the application due to a specific exception that applied to that case. Simply correcting the party names and paragraph numbers would not rectify the error. The reviewer must limit the scope of the proposition, determine if the exception applies to his/her client's situation and possibly re-evaluate the organization of the pleading. Therefore, citation checking cannot be outsourced to another generated answer that only restates the original generated answer's understanding.

Unsupported propositions and missing adverse material

Verification should be organised around a proposition ledger. Each material statement is linked to its supporting authority or evidential record, the role it plays in the argument, and the person who checked it. Unsupported statements are not all of the same kind. Some concern missing evidence, some concern a disputed interpretation, and some are recommendations based on professional judgment. These should be labelled differently. A recommendation can be defensible without being a holding, provided it is not presented as established law.

Omissions require a separate search strategy because they are not visible in the generated draft. The reviewer should test competing interpretations, possible exceptions, contrary factual accounts, and material adverse authorities. The exact duty to disclose adverse authority to a court depends on the applicable professional and procedural rules; this chapter does not assume identical obligations throughout the sample. As a quality-control matter, however, counsel needs to know the strongest relevant contrary material in order to assess the advice and avoid an incomplete presentation.

An adversarial verification pass therefore asks a different question from the drafting pass. Rather than requesting further support for the chosen conclusion, it asks what could make that conclusion inapplicable. The search should include different terminology and sources outside the initial result set. A restricted collection may be appropriate for the task, but its boundaries must be recorded. The absence of an adverse decision from that collection does not establish that no adverse decision exists. Where important sources are inaccessible, the advice should explain the resulting limitation.

General-purpose and selected-source systems

A general-purpose system and a system connected to selected legal sources differ in the material they can use and the ways in which that material is presented. Those differences should be assessed rather than converted into an assumption that one category is invariably reliable. A selected collection may improve traceability, but an answer can still misstate a retrieved passage or overlook an authority outside the collection. A general-purpose tool may be useful for reorganising a verified outline while remaining unsuitable as the unexamined source of its legal propositions.

Magesh and colleagues distinguish correctness from whether an answer is supported by its cited authority. Their historical United States evaluation tested three retrieval-based products and found persistent errors despite retrieval (Magesh et al., 2025, §§4-6, pp. 220-225). This chapter uses the distinction to structure verification, not to import a foreign error rate into local practice or rank present products. Product coverage, test date, query selection and the distinction between an individual answer and a complete professional workflow limit the inference. A local evaluation must establish its own tasks, relevant sources, language coverage and release conditions.

Procurement and practice decisions should accordingly separate collection quality, retrieval quality, and answer quality. Collection review asks what sources and periods are covered. Retrieval review asks whether relevant material can be located. Answer review asks whether the output accurately uses that material for the particular question. Local-language coverage, unreported decisions, and specialised procedural materials should be investigated where relevant, rather than inferred from the size of a database. A system can perform well at one stage and poorly at another.

From a benchmark result to a local verification design

A legal benchmark and a professional release decision answer different questions. A benchmark can reveal recurrent failure types under a specified test; a release decision concerns whether this document may be used for this client or court. The two should be connected through an error taxonomy. For each locally consequential proposition, the evaluator should distinguish missing authority, incorrect authority, an unsupported inference, an omitted qualification, a wrong applicable date and a mismatch between the question and the answer. Recording only whether the final answer sounds correct collapses defects requiring different remedies.

The proposed local test uses a reserved set of questions assembled by practitioners before testing the selected configuration. It includes ordinary work and deliberately difficult cases drawn from the institution's scope: an amended rule, an adverse authority, a disputed date, a language transition and a request with a false premise. A difficult-case set diagnoses failure; it does not estimate how frequently the failure occurs in ordinary work. A separate sample of routine tasks is needed for that estimate. Reviewers should

record disagreements and their resolution, so that the benchmark does not conceal uncertain law behind a single answer key.

The comparison should also include abstention. A tool that supplies fewer answers but correctly identifies an unresolved legal issue may be preferable for a consequential task to one that always completes the form. Yet referral is not costless: the relevant comparison includes the additional professional time and whether a person reaches competent assistance. This yields a two-stage assessment: first determine whether released work satisfies the legal quality threshold; then compare the resources required to reach it. A time saving produced by shifting an unresolved defect to a client does not pass the first stage.

Pleadings, contracts, and document review require different checks

Pleading review must connect factual allegations to instructions and evidence, not merely to plausible narrative. Dates, amounts, party identities, procedural history, requested relief, and admissions require particular care. A generated paragraph should not turn an unresolved inference into an assertion of fact or introduce a concession because it makes the argument more coherent. The responsible lawyer should also confirm that the document responds to the actual procedural stage and applicable directions. A formally polished pleading is not necessarily a procedurally appropriate one.

Contract drafting requires a different form of source control. An approved precedent can organise the work, but the reviewer must reconcile defined terms, commercial instructions, schedules, cross-references, responsibility allocations, and termination arrangements. A change in one clause may alter the effect of another. The use of a familiar clause is not evidence that the client has authorised the risk it contains. The release record should identify material departures from instructions or precedent and the person who approved them. Legal enforceability still requires the governing law to be examined.

Document review should preserve the relationship between each extracted statement and its location in the original file. Classification for relevance, privilege, confidentiality, chronology, and disclosure can involve different tests and different reviewing roles. For a large collection, the team should retain an inventory of received material, excluded or unreadable files, duplicates, attachments, and version relationships. Quality testing should examine material selected by the system and material it left out. Otherwise the review can detect false inclusions while remaining unaware of consequential omissions.

A hypothetical disclosure exercise illustrates the issue. A model marks an email as routine and omits its attachment, although the attachment contains the disputed instruction. Sampling only the documents marked relevant would not test this failure. The team needs a method for inspecting excluded material and incomplete document families, with targeted review where the legal consequence is serious. The sampling

design and escalation threshold are task-specific judgments; no universal sample size or acceptable miss rate is supplied here.

Confidentiality belongs before the first prompt

Output checking cannot undo an inappropriate disclosure made during input. Before client material enters a system, the team should establish what information is necessary, the approved processing environment, who can access it, how it is retained, and whether it can be reused. These questions concern both the selected system and the particular account configuration. A promise about a product category is not sufficient evidence of the conditions applying to an individual user's session. The relevant confidentiality, privilege, data, contractual, and court restrictions must be identified separately.

Removing names is not automatically enough: a detailed factual combination may still identify a person or disclose a confidential strategy. The practical choice may be to use an abstracted problem, an approved internal environment, or a non-AI method. Authorisation to use a tool does not mean that every record can be entered into it. The chapter's proposed policy requires a designated reviewer to resolve uncertain input permissions before processing, rather than leave a junior user to infer permission from the availability of the upload button.

The time saving must survive verification

Verification intensity should reflect the consequence of error, difficulty of independent checking, novelty of the task and opacity of the supporting process. These are reasons for allocating review effort, not measurements of efficiency. The appropriate comparison is between complete workflows producing work of the same acceptable quality. Prompt preparation, source retrieval, factual checking, correction, supervisory review and post-release remediation belong in that comparison. Time to the first draft is only one component.

An explicit accounting identity helps prevent misleading claims: net time saved equals the time required by the credible non-AI workflow minus the total time required by the AI-assisted workflow, including its checks and corrections. This is an analytical definition, not a calculated result. Separate records should identify who performs the work. A reduction in senior drafting time accompanied by additional junior verification may redistribute cost rather than eliminate it. Time saved is also not evidence that a legally necessary check can be omitted.

Checking outweighs the claimed saving when the controlled workflow takes longer than a realistic alternative, or when the purported saving depends on leaving a consequential uncertainty unresolved. The answer can differ across tasks within the same matter. An approved clause comparison may save time while reconstructing a generated chain of authorities does not. The appropriate response is to narrow or discontin-

ue the unhelpful use, not reduce scrutiny until the efficiency claim becomes true. Evaluation should therefore compare task-specific release-ready work, not count words generated.

A pre-release review protocol

Table 2.1 turns these distinctions into a release record. The drafter assembles the evidence; a reviewer with the relevant competence accepts or rejects the work; and the person responsible for the client communication or filing approves the final version. Small practices may combine roles, but should not describe a self-check as independent review. The protocol is an institutional proposal requiring adaptation to the governing professional and procedural rules.

Table 2.1. Pre-release protocol for AI-assisted legal work

Review stage	Check and retained evidence	Responsible reviewer and release condition
Task and input permission	Record purpose, recipient, authorised data, approved environment, relevant restrictions, and the intended role of the output.	Matter owner approves use. Unresolved confidentiality or scope questions are referred before processing.
Authority identity	Open the original authority; retain title, issuing body, date, identifier, pinpoint, and source location.	Researcher assembles the record; legal reviewer rejects unsupported or untraceable reliance.
Applicable version and status	Check relevant time, amendments, commencement, subsequent treatment, jurisdiction, hierarchy, and procedural context.	Legal reviewer confirms the status in which the source is used; material uncertainty is communicated or resolved.
Holding and interpretation	Read the surrounding reasoning and qualifications; distinguish holding, submission, quotation, example, and inference.	Legal reviewer explains why the passage supports this proposition in this matter.
Adverse and omitted material	Test competing interpretations, exceptions, contrary authorities, collection boundaries, and missing evidence.	Reviewer records the search and handles any disclosure duty under the applicable rules.
Factual and client instructions	Check allegations, identities, dates, amounts, procedural history, admissions, relief, and material drafting changes.	Responsible professional resolves discrepancies with the record or client before release.

Table 2.1. Pre-release protocol for AI-assisted legal work

Review stage	Check and retained evidence	Responsible reviewer and release condition
Contract or review-specific checks	For contracts: reconcile definitions, schedules, risk allocations, and approvals. For document review: test exclusions, attachments, and coverage.	Task specialist records checks and unresolved gaps; consequential gaps prevent ordinary approval.
Language and procedure	Check consequential wording, translation status, form, filing requirements, service steps, and the actual deadline basis.	Qualified reviewer confirms the task-specific requirements; no unverified certification or filing status is represented.
Final version and communication	Retain the approved version, reviewer, approval date, limitations, recipient, and applicable disclosure decision.	Release owner confirms that the delivered work is the reviewed work and appropriate for the stated use.
Post-release error	Preserve the affected version; identify reliance; assess notice, correction, procedural action, and related affected work.	Matter owner coordinates the competent correction route and records closure and remaining consequences.

Local professional rules, court directions, privilege, retention requirements and client arrangements govern completion of this protocol.

A document should leave this process with a clear status: approved for the stated use, approved with an expressly communicated limitation, or withheld pending correction. Approval attaches to a version and purpose. Later changes to a citation, an allegation, a contractual term, or the requested relief may require renewed review. After release, discovery of a material error should trigger preservation of the relevant record, assessment of who relied on it, and correction through the applicable client or court procedure. Verification is complete only for the work actually delivered, not an earlier draft that happened to be sound.

2.2 Court Administration and Case Management

Administration determines the conditions of a hearing

Court administration does not become legally insignificant because it precedes adjudication. Filing categories, hearing dates, assignment decisions, and transcript records shape what can be considered and when. A recommendation that remains visible to an authorised officer may be readily corrected. The same recommendation becomes more consequential when it determines a queue, prevents a filing from progressing, or reaches the judge as an apparently settled record. Analysis should therefore follow the administrative output to its practical effect, including effects that occur before any formal decision is issued.

Registry work requires its own allocation of power. A person authorised to correct a database field may lack authority to excuse late filing, transfer a case or vary a judicial direction. The protocol therefore separates technical rectification from procedural relief. Its legal foundation must be the court's own rules and lawful directions, read against hearing and participation rights. This distinction prevents a useful administrative remedy from being mistaken for the remedy needed to restore a lost procedural opportunity.

The distinction is visible in South Africa’s constitutional text. Section 34 protects access to a fair public hearing before a court or appropriate independent and impartial forum; section 35(3)(d) addresses unreasonable delay in criminal trials. These are constitutional provisions, not AI-specific permissions or detailed scheduling rules (Constitution of the Republic of South Africa, 1996, ss. 34, 35(3)(d)). Their relevance to this analysis is that an administrative process must be examined in relation to the hearing it enables. A more efficient diary is not the whole legal objective.

Digitised filing and inferential classification

Digitised filing changes the medium of submission and storage. It may include ordinary programmed checks, such as whether a required field has been completed. AI-based classification adds an inferential task: the system interprets a document or narrative and assigns it to a category. These functions can appear in the same portal but should be assessed separately. A system may receive a document successfully while classifying it incorrectly; it may classify it accurately while failing to preserve an attachment. Neither event is described adequately by the statement that electronic filing is operational.

The filing record should distinguish receipt, technical validation, administrative acceptance, and any decision about legal admissibility. A technical error message should not silently acquire the status of a judicial refusal. Equally, a receipt should not be described as confirmation that every applicable filing condition has been met. The person submitting the material needs to know which event has occurred, which remains

outstanding, and how an unresolved issue can be brought to the responsible officer. The legal effect of each event must come from the applicable procedure.

Classification should remain reviewable at the point where it changes treatment. A proposed category such as urgent, incomplete, duplicate, or incorrectly filed needs a documented definition and a record of the consequence attached to it. The interface should permit uncertainty rather than require every matter to fit a confident label. For consequential exceptions, a reviewer needs the original filing, not only the system's summary. An applicant's explanation should remain available even when it does not fit the category on which the initial recommendation was based.

A hypothetical application for urgent relief demonstrates why timing matters. The text primarily describes a contractual dispute but also identifies an imminent event that may make later relief ineffective. A classifier routes it to the ordinary list because its topic resembles routine commercial litigation. The substantive concern is not simply that the label is wrong. It is that the proposed relief may lose practical value while the matter waits. The protocol should permit an urgent human review without requiring the applicant first to prove that the model made a technical mistake.

Scheduling, prioritisation, and workload measures

A scheduling process will include non-interchangeable constraints. Constraints on judicial availability, procedural timetables applicable in the matter, whether one or both parties participate in the matter, who represents the parties in the matter, whether an interpreter is required, how many witnesses there are (and their schedules), where the hearing takes place and whether immediate consideration of the matter is necessary may all impact what constitutes a lawful hearing. A system optimized solely to fill available slots could produce a schedule that fails these constraints. The scheduling owner must identify which constraints are mandatory, which represent a preference, which are discretionary and which require input from a judge. If a constraint has not been resolved it should appear to have not been resolved instead of having appeared to have been resolved when actually it was not.

When evaluating priorities, specific attention must be given to the objective that is being optimized. While some courts may focus upon closing the number of matters, others may focus upon the age of the pending matters or the time elapsed until first hearing or the time elapsed until effective relief. Optimization based upon readily completed matters can result in a reported backlog being reduced while not resolving more complex matters. Optimization based exclusively upon age may overlook new, recently filed and urgent matters. These are examples of optimization methods to consider, they do not represent any reported results of a particular court. Priorities should be evaluated to determine why the chosen method reflects the legal function of the court and which waiting times are excluded.

Resource utilization should differentiate between case counts and the amount of work required to resolve a case. The number of documents within a file does not always equate with the complexity of the case and the duration of a hearing does not always reflect the level of importance associated with that hearing. Resource allocation forecasts can utilize estimated values however those estimated values should not evolve into unreasonable expectations that obligate officers or judges to perform less work than reasonably necessary to effectively resolve a case. Thus, the proposed protocol provides for a distinction between administrative resource planning forecast estimates and instructions related to the resolution or disposition of a single case.

In addition, external resources utilized by parties involved in the litigation should be included when evaluating performance. A rescheduled hearing may allow the court to avoid additional administrative expense while imposing yet another travel requirement or additional absence from employment or additional need for an interpreter. Further, although a digital notice may be technically transmitted it may still fail to be either read or received by its intended recipient. All of these requirements necessitate evidence unique to the institution. As stated above this chapter makes no claims regarding the size of a backlog that can be eliminated or the improvement in service level that can be achieved through implementation of this protocol. This chapter assumes that prior to making such a claim, the court must specifically identify the population upon which it intends to make such a comparison, over what comparative time period, excluding whom and what is required to correct such inaccuracies.

Case allocation and judicial assignment

Subject Matter Routing vs. Assignment

Subject matter routing recommendations cannot simply be viewed as a one-step technical exercise; they are distinct from rosters which are further apart from decisions regarding recusal or transfers. Each level of personnel involved in the allocations may possess varying degrees of authority. The allocation record should include details on the underlying framework, classification used (if any), restrictions involved and the individual(s) empowered to approve / reverse the allocated designation. A vendor's capacity to modify their own routing algorithms has no bearing on their authorization to allocate judges to cases.

Automation enables reproducibility of an allocation process however does not ensure lawfulness or independence. Although an algorithmic rule may be transparent, it may still be predicated on an inappropriate parameter. An inferential recommendation may be useful while remaining difficult to reconstruct; that possibility needs examination in the particular system. The reviewer should determine if the allocation criteria were authorized, if there are controls over exceptions, and if potential stakeholders can act in contravention of the applicable rules during the course of the process. A simple log of the assigned designation is insufficient when the key issue is why another eligible designation was precluded.

In terms of corrective action, the distinction between a clerical modification and a procedural instruction must be preserved. The person responsible for maintaining the data base may be capable of modifying a wrongly coded entry; however, he/she would not have sufficient authority to relocate a case or alter a judicial order. To accomplish this, the proposed work flow will require a defined escalation path to the officer/judge who has been delegated sufficient authority and/or access. Affected parties should be provided reasonable notice and opportunity to respond consistent with the governing procedures. "Contact Technical Support" is clearly insufficient response to a challenge related to the legitimacy of a given assignment.

Transcription and Administrative Record

Transcription allows hearings / communications to become records that can subsequently rely upon. Therefore, quality assessments should focus on legally significant meaning (not simply overall word accuracy). Omission of a negative statement, misrepresentation of an amount, confusion by a speaker, substitution of certainty for uncertainty may be important regardless of readability of the transcript. Identification of the status of the transcript (working aid / provisional record / certified record/part of official proceeding) should occur through reference to applicable rules / practices of the institution.

The recommended recordkeeping practices maintain an acceptable linkage back to the original recording which was permitted to be made and recorded as such. The methods provide identification of unclear sections / speaker transitions / modifications / corrections / approval of the operative document. Edited versions of transcripts should not hide the fact that a critical modification was made. When a party challenges the record, it should be clear who has access to the source document and how a revised version is transmitted to entities that have previously relied upon the previous version. The technology service company could assist in reconstructing documents without taking a position regarding the legal dispute over content.

Additionally, transcription and translation need to continue to be distinguished from each other. Translated transcriptions involve at least two conversions and defects can emanate from either. A translator may faithfully translate an already erroneous transcript. Thus, disagreements regarding subsequent wording should refer back to the original permitted recording wherever possible utilizing language assistance appropriate thereto. Fluency alone does not establish compliance with the applicable certification or interpretation requirements.

Correction burdens, notification, and continuity

Correcting errors also involves timing. Changing a priority code on a current document is changing one existing record; it doesn't bring back the opportunity to have heard the case since it was misdirected due to the error. Therefore, you will use separate analyses for the actions of amending, notifying, reconsidering, and restoring. These are the authorially analytic categories as they reflect that there could be some intervening consequential activity once an error and its existence become known. The institutional procedure establishes which corrective activities can be performed by the institution.

The notification process must relate to the intervening consequential occurrence and the responsive option. For example, an individual needs a clear definition of what has occurred, the new date or condition that is operative, and how he/she/it can query it. A general statement indicating that a particular system was used provides minimal value. If the institution is unable to publicly disclose the detail regarding an intervening consequential occurrence, then the institution must indicate how a relevant explanation or document may lawfully be made available. The procedures establishing who will receive notifications (i.e., the recipient(s)), how notifications shall be delivered (e.g., electronic mail, phone), and within what time frame such must occur, must be established by the institutional procedures not solely by the software systems being utilized.

An operational exception pathway must include: (1) ownership; (2) an entry point into the exception pathway; (3) preservation of a unique identification number for each case filed via the exception pathway; and (4) a means to communicate directly with decision-making authority prior to delay defeating the reason for the exception pathway. An individual submitting issues through this path should not have to know if the problem arose in the classification process, data base, calendar function or messaging service. However, internally the court must know where problems exist so that systemic corrections can be made to prevent future problems.

Corrective action involving a single instance and corrective action involving systemic problems are two distinct processes. Completing an incident report is not evidence that either has been accomplished.

Planning for service continuity must focus on institutional legal obligations rather than simply planning for server uptime. In addition to mechanisms for providing notice during outages, there must be procedures for accepting urgent communications, recording attempted submissions, maintaining timelines for transactions and decisions, identifying pending cases, and communicating new arrangements. The determination of whether a deadline has been extended or an alternate submission mechanism is valid requires competent authority decision making under applicable rules. Institutional policies must not imply that compliance with such deadlines automatically insures continued validity of an otherwise invalid transaction.

Restoration creates its own risks. Material submitted through an alternative communication pathway must be processed against a restored system without creating duplicate records, losing records, or disrupting the order of events. Incident reporting documents should identify: (a) the duration of the time when services were interrupted; (b) all affected transactions; (c) all affected decisions; (d) all affected notices; and (e) all corrective action taken. There must also be staff training and alternatives for accessing critical information during periods of non-availability of normal pathways of access, i.e., a plan for finding a pressing matter when the primary database is down. An emergency plan that has never been practised does not demonstrate operational continuity.

A consequential-administration workflow

The strongest case for automation is that existing administrative processes may themselves involve delay, inconsistent classification, and burdensome manual handling. The answer is not to preserve them unchanged. It is to compare the proposed workflow with a documented and realistically improvable alternative. Better indexing, clearer forms, staffing adjustments, and ordinary rules-based routing should be considered alongside inferential tools. The appropriate design may combine them, reserving AI for a narrow task and preserving direct access to an officer for matters outside that task.

Table 2.2 identifies the minimum record at each consequential step. "Registry officer," "listing officer," and similar labels denote functions to be assigned locally; they are not claims that every jurisdiction uses those titles or grants the same powers. Before use, the court should replace each role label with the competent office, cite the applicable source of authority, and specify the actual notification and escalation arrangements.

Table 2.2. Court-administration workflow for consequential steps

Step and consequence	Responsible function	Notification	Exception or correction route	Audit record
Receipt and validation: material may not enter the case record.	Registry intake authority.	Confirm receipt status, defects, and outstanding steps.	Preserve attempted submission; refer disputed legal effect to competent authority.	Timestamp, files, defect, receipt, officer action.
Classification: category changes handling or access.	Designated classification reviewer.	Explain the consequential category and query route.	Allow explanation outside standard fields and examination of original material.	Input, category, model/rule version, review and reason.

Table 2.2. Court-administration workflow for consequential steps

Step and consequence	Responsible function	Notification	Exception or correction route	Audit record
Urgency: delay may defeat effective consideration.	Authorised urgent-list officer or judge.	Acknowledge the urgency request and immediate next step.	Direct human escalation independent of a negative automated flag.	Urgency facts, referral time, responsible actor, direction.
Assignment: court, list, or judge is selected.	Officer or judge authorised by the allocation arrangement.	Give required assignment notice and relevant procedural information.	Separate clerical amendment from transfer, recusal, or judicial direction.	Governing arrangement, criteria, exclusions, approval, change history.
Scheduling: hearing or response opportunity changes.	Listing authority, with judicial direction where required.	Communicate date, conditions, changes, and query procedure through authorised channels.	Address representation, interpretation, participation, and mandatory constraints.	Constraints, proposed and confirmed dates, notices, corrections.
Transcription: text enters the operative record.	Record officer and authorised approving function.	Identify provisional or approved status and dispute route.	Refer material wording or speaker disputes to authorised source review.	Source link where lawful, uncertainty, versions, correction and approval.
Outage: ordinary receipt or access is unavailable.	Continuity owner and competent procedural authority.	Publish usable alternative arrangements and limits.	Use approved urgent channel; obtain any necessary deadline or procedural direction.	Affected period, attempted acts, alternative receipts, incident decisions.
Recovery and downstream correction: prior consequences may persist.	Registry correction owner, coordinated with the deciding authority.	Notify relevant parties and recipients of operative changes.	Reconcile records; reissue notice or seek reconsideration where necessary.	Affected cases, amendments, notifications, directions and completion checks.

Role labels identify functions to be assigned locally, not universal offices or powers. Every consequential step requires its actual source of authority and notification procedure.

Administrative recommendations affect substantive interests when they materially alter access, timing, assignment, the record available for decision, or the practical usefulness of relief. The person who can correct the problem is not necessarily the person who operates the software. A reliable court-use protocol makes that difference explicit, preserves the evidence needed for correction, and connects a technical amendment to the procedural action required to put the matter back on a lawful course.

2.3 Judicial Reasoning and AI-Assisted Decision-Making

The boundary concerns judgment, not authorship alone

The boundary between assistance and delegation cannot be drawn simply by counting how many words a judge writes personally. Research, organisation, translation, and drafting support can all contribute to judicial work. The more important question is whether the judge remains responsible for establishing the relevant facts, identifying and applying the governing law, addressing the parties' positions, and deciding the matter. A human signature on a generated document does not answer that question; nor does the mere use of software establish that judgment has been displaced.

Judicial assistance should be examined at the point where an output acquires legal weight. Research, record organisation, inference, reasons and disposition engage different responsibilities. The following analysis first identifies the decision-maker's retained function, then compares enacted judicial-governance instruments. It treats transparency and human control as distinct: an openly disclosed output may remain unverified, while an independently checked output may still require disclosure under the governing rule.

Colombia supplies a concrete judicial authority. In *Sentencia T-323/24*, dated 2 August 2024, the Constitutional Court distinguished administrative, documentary, corrective, and summarising assistance from substitution for non-delegable human reasoning concerning facts, evidence, the justification of a decision, and its adoption. It linked judicial use to transparency, responsibility, verification, and privacy, while describing its guidance as an initial approach requiring attention to later legal and technical developments and the particular jurisdiction (Corte Constitucional de Colombia, 2024, paras. 295-300). This is a dated Colombian holding, not a rule automatically applicable elsewhere.

Research assistance and the evidential record

Research assistance is most readily separated from adjudication when the output supplies identifiable material that the judge can examine independently. A list of authorities, a proposed chronology, or a cross-reference between submissions can support work without settling the contested issue. Its usefulness depends on the judge's ability to return to the sources and identify omissions. A summary that becomes the only accessible version of the record occupies a different position: it controls what is available for judgment, even if it is formally described as assistance.

The proposed protocol therefore distinguishes a research output from a finding. The former identifies what a document states or where an issue appears. The latter accepts, rejects, or assigns significance to that material. A generated chronology should preserve uncertainty and competing dates; a summary of submissions should not convert an allegation into an established event. The reviewer should identify the source of each material entry and maintain the distinction between evidence, argument, inference, and the court's conclusion.

The evidential record also defines an important boundary for outside information. A model-generated factual statement about an event, person, practice, or technical matter is not made part of the case merely because it appears in a research response. The court must identify whether and how the information can properly be considered under the applicable procedure. Where participation or challenge is required, it should occur before material reliance, not after the statement has silently entered the reasons. The protocol does not create an exception to ordinary evidential requirements.

A hypothetical example illustrates the difference. A system compares two witness accounts and notes inconsistent times. That comparison can be checked against the record. A further output states that one witness is dishonest because the inconsistency resembles a pattern in other cases. The second statement adds an evaluative inference requiring its own legal and evidential basis. The court cannot validate it merely by verifying that the two times differ. The inferential step, and not only the underlying extraction, is the object of scrutiny.

Factual interpretation and credibility assessment

Fact assessment requires attention to relevance, context, competing explanations, and the applicable burden and standard of proof. An output that predicts which account resembles previously accepted testimony does not, without more, establish whether that account should be accepted in the present case. Historical acceptance and legal correctness are distinct targets. The task-based protocol should therefore state what the system estimates and prevent an estimate from being presented as proof of the proposition the court must decide.

Credibility assessment deserves separate treatment because a score can conceal a chain of contested assumptions. A proposed tool may draw upon wording, omissions,

response time, tone, or other features. The assessor should ask whether those features are legally relevant, whether the method is valid for the setting and population, and whether competing explanations can be examined. A score should not be treated as a direct observation of honesty. This is an analytical limit on inference, not a finding that any particular product uses those features or that a uniform statutory prohibition exists.

For institutional policy, this chapter proposes excluding automated credibility determinations from ordinary assistive use. Any contemplated exceptional use would require a separately established legal basis, a defensible evidential role, appropriate validation, and a real opportunity for challenge. The restriction is proposed because the consequences and assumptions differ materially from source retrieval. It does not imply that human credibility assessment is infallible. Rather, replacing one uncertain process with another requires evidence about the replacement and its reviewability, not confidence in the label "objective."

Bail, sentencing, and other consequential recommendations

Combining bail and sentencing in a single, generic prediction task would make little sense. Each has fundamentally different legal questions, relevant factors, decision makers and procedural safeguards. To assume that a model estimates an event means that the model is estimating a criterion that a court is permitted by law to use would be entirely wrong. Regardless of how the targeted outcome relates to the model's output, the judge still needs to determine how the model's output relates to his/her own judgment as to what consequence follows from that fact. That relationship cannot be determined based on a vendor's claims regarding their model's accuracy nor will it be clarified by simply adopting a foreign court's application of another system.

The proposed assessment process first identifies the target of interest, data sources used, population assessed, validation timeframe, degree of uncertainty, threshold value(s) and institutional usage of the assessment. The next step in the process is to identify whether the appropriate law allows the judge to utilize the estimate, provide an explanation for the estimate if so allowed and allow the defendant to dispute errors within the inputs and/or flawed assumptions in those inputs. A positive response to one inquiry does not ensure a positive response to either of the other two inquiries. For example, while a transparent estimate may provide a basis for relevance it may nonetheless fail to provide a sufficient level of validation; and even though a validated recommendation may be made, it may still fail to be utilized due to lack of ability to establish requisite procedural safeguards.

Additionally, meaningful control will require a legitimate ability to deviate from a recommended course of action. As previously discussed, judges must be capable of reviewing material outside the scope of the model and explaining their ultimate disposition through legal criteria established by statute or case law. Thus, an interface that requires a justification only when the judge rejects the model's recommendation may

discourage departure without explaining how the output was produced. Therefore, whether such an interface actually exists will depend upon observation as opposed to assumptions. The institutional review should therefore assess defaults, potential impact on the judge's workload, approval requirements and a record of how recommendations entered into actual judgments.

In this regard, the key difference lies between providing assistance that remains subject to independent evaluation and providing an artificial intelligence-based computational recommendation that substitutes for a legally assigned function of judgment. By placing bail and sentencing recommendations outside normal permitting processes under this model policy, the classification is intentionally conservative. The liberty implications associated with bail recommendations and the nature of contested predictions along with the inherent difficulties associated with reversing reliance justify a separate determination on authority and validation. Such classification does not imply that all six focal jurisdictions have enacted identical prohibitions.

Preparing Reasons and Preserving Independent Judgment

Preparation of reasons involves more than preparing readable prose. Preparation of reasons links the resulting disposition to: (1) accepted facts; (2) applicable law; and (3) issues requiring resolution. Assistance via artificial intelligence should therefore be evaluated relative to which component of that preparation process it undertakes. Editing an independently developed rationale differs significantly from developing conclusions and generating rationales that a judge then adopts. The same interface can support both formats. Accordingly, the assessment should describe the workflow employed and not categorize the end-product as having been created for all purposes.

One effective means of ensuring that judges maintain independent control is to preserve the sequence of work at points where it matters: (1) identification of issues; (2) materials reviewed; (3) judicial assessment; (4) assistance sought; and (5) final modifications adopted. Preserving such documentation need not result in wholesale disclosure of every note or draft document. Material preserved pursuant to lawful methods must protect privacy rights in relation to judicial deliberations and personal information. The goal is to preserve sufficient documentation to differentiate between assistance involving expression versus an unreviewed transference of reasoning.

Judges' independent rejection of outputs must be possible both conceptually and practically. Conceptual independence requires an understanding of the particular proposition being made by the model, including the boundaries thereof. Practical independence requires: (1) access to source documents; (2) adequate time to evaluate those documents; (3) authority to reject recommendations; and (4) a viable alternative when the system fails. Hence, training for judges using artificial intelligence models; access to records related thereto; and protected review times are essential conditions for maintaining independence.

Additional burdens imposed by the need for verification can further exacerbate existing burdens. In paragraph 362 of T-323/24, the Colombian Court explicitly noted that utilizing a general-purpose AI tool to search for specific information followed by verification of those sources could potentially increase rather than decrease judicial workload (Corte Constitucional de Colombia, 2024). While this comment represents no measured assessment of efficiency costs, it underscores an important distinction made throughout this chapter: differences in rate of generation as compared to time required for responsible utilization of AI-generated outputs.

Disclosure and participation

Material Contribution

The appropriate time to consider notice prior to relying on AI-generated outputs depends on the type of material introduced by AI. Where an AI output introduces a factual assessment (e.g., a new finding), a contested legal proposition, an evidentiary classification, or a consequential procedural recommendation that has not been previously addressed by the parties, then material contribution should trigger consideration of notice before reliance. The notice should indicate:

- the nature of the task;
- what materials were used;
- relevant limitations; and
- available routes for response.

Notice should not require a party to fully comprehend the model to raise objections. Objections may involve missing facts, incorrect authorities, inappropriate uses, etc. Rather than objecting to the architectural design of the model itself, objections are likely to be based on some other aspect of the output (e.g., how it was generated).

Spelling Corrections vs. Evidentiary Assessments

While many corrections made by AI are minor (spelling corrections, grammatical adjustments) and would not normally require formal notice and/or a challenge route, others are more significant (e.g., changes to a party's pleading). Some jurisdictions may require such significant changes to occur through specific procedures (e.g., motions to amend).

In addition, even where formal notice and/or a challenge route is provided, excessive, undifferentiated disclosure can obscure the contributions that actually have an impact and could result in protected materials being disclosed inappropriately. The proposed policy seeks to establish a functional distinction between required notice and discretionary disclosure. Depending on applicable laws, however, notice may need to be broader or take a different form.

Documentation and Confidentiality Concerns

Documented decisions regarding disclosure should be made by the judicial official(s) who utilize AI systems. Judicial officials should be aware of potential concerns relating to confidentiality, deliberative process protections, equal participation rights, and possible impacts on the resolution of the dispute. Documentation also supports the ability to assess whether reliance on an apparently administrative contribution changed the course of the case.

Challenge Route

A challenge route should lead to a person with authority to resolve the dispute. Possible responses to a challenge may include reviewing the source data used to create the AI output, amending the record if necessary, excluding unreliable input(s), requesting supplemental submissions from the parties, and/or reconsidering a procedural step in light of applicable law. Technology suppliers will typically be best positioned to explain technical details underlying their products' operations. Suppliers, however, should not determine legal sufficiency of a challenge.

Trust Not Substituted for Disclosure

Reliance cannot be cured by instructing a party to "trust" the system when there exists a lack of disclosure regarding the basis for that reliance.

Table 2.3. Task-based protocol for judicial AI assistance

Task and proposed policy status	Required control	Disclosure, escalation, and retained record
Public-source research: potentially assistive within approved scope.	Judge or qualified assistant examines originals, relevance, currency, and contrary material.	Record sources and the role of the output. Apply governing disclosure rules and assess any material new proposition.
Chronology or record summary: assistive only with source access.	Preserve contested facts, uncertainty, omissions, and links to the underlying record.	Escalate untraceable or disputed material; do not promote a summary into a finding without judicial assessment.
Factual interpretation: requires distinct legal and evidential assessment.	Identify the inference, permissible evidential role, validation limits, and independent judicial evaluation.	Consider party participation before material reliance. Retain the relevant basis and review decision under lawful access controls.
Credibility assessment: excluded from ordinary permission under this model policy.	Do not treat a score or linguistic pattern as a direct finding of truthfulness.	Any exceptional proposal requires separately established authority, defensible evidence, validation, and challenge; local prohibitions control.

Table 2.3. Task-based protocol for judicial AI assistance

Task and proposed policy status	Required control	Disclosure, escalation, and retained record
Bail or sentencing recommendation: no routine authorisation inferred.	Identify the precise legal question, permitted factors, model target, population, uncertainty, and power to depart.	Require use-specific authority and safeguards before reliance. Record explanation, party challenge, judicial assessment, and unresolved limits.
Drafting reasons: distinguish editing from substantive judgment.	Judge retains assessment of facts, law, issues, and result; verify every adopted material addition.	Preserve an appropriate record of assistance and changes, respecting deliberative protection; assess material disclosure duties.
Automated adoption or execution of an outcome: outside ordinary assistive approval.	Identify the legal allocation of the function and any valid authority for the contemplated arrangement.	Do not proceed through a label or signature alone. A prohibited delegation remains prohibited despite the table's safeguards.

The Colombia-Brazil comparison below applies the permission, human-control and disclosure questions to actual instruments. A safeguard cannot supply missing legal authority.

Meaningful judicial control requires more than the final ability to click "approve." It requires independent command of the legal and evidential basis, practical freedom to reject the output, and a procedure through which material reliance can be tested when the law requires it. The resulting decision must be attributable to the judge's exercise of the assigned function. Assistance may reduce preparation work; it should not make the authority, content, or contestability of judicial judgment disappear into the system that prepared the draft.

A matched comparison of judicial summarisation in Colombia and Brazil

The tasks compared share a specific objective: to evaluate the performance of a new process (a judicial officer using AI to summarize documents and provide a draft of a preliminary explanation of their findings) where the judge has access to the full file. The comparison includes Colombia and Brazil, since both countries have enacted specific judicial regulations regarding AI systems intended to assist judges, allow them to monitor how such systems work and maintain openness and transparency. It is not about comparing the level of technological advancement of each country. Rather, the comparison seeks to find out if the same process meets the obligations of both regulations and if there are any aspects that would remain unchanged, despite a common commitment to accountability in the use of AI.

Colombia's Agreement PCSJA24-12243, adopted on 16 December 2024 following T-323/24, sets judicial AI guidelines. Articles 4-5 govern permitted uses, while articles 6-9 address suitability, knowledge, duties and verification. Article 8(3) excludes AI assessment of evidence and resolution of legal problems; article 8(4) prohibits applying law, giving reasons or deciding a case exclusively from a tool's responses. Article 10 requires express disclosure of generative AI used for institutional work. Particularly where use affects substantive decisions, the record must identify the tool, model, provider and use date; acknowledge review and acceptance of its terms; explain purpose, manner of use and prompts; and locate and distinguish incorporated output. Disclosure may appear in the main record, footnotes or annexes (Consejo Superior de la Judicatura, 2024, arts. 4-10).

Brazil's CNJ Resolution 615/2025 classifies supervised summaries and judicial drafting support as low-risk uses in Annex BR1 and BR4, subject to their stated human-control conditions. Assessment and valuation of evidence in contested proceedings fall within high-risk category AR2. Calling a task a summary does not settle its classification. Under article 19(6), a judge may mention generative drafting assistance in the decision, but automatic recording in the court's internal system is required. Article 21(2) separately requires AI use in automatically generated products to be identified in system logs through adequate, comprehensible labels. Public mention, internal traceability and any procedural duty to inform parties therefore require separate analysis (CNJ, 2025, arts. 19(6), 21(2), Annex BR1, BR4, AR2).

When applied to this joint task, the first question is whether a summary provides an accurate reflection of the parties' submissions as submissions. A submission that a party alleges payment has been made is distinguishable from a finding that payment has been made. The second question is whether the proposed draft creates an inference not present in the judge's assessment. Even though approval is granted, these are content-based inquiries. The third focuses on routes of deployment and relevant restrictions on external tools. The fourth examines information recorded and disclosed. A single declaration that the judge reviewed the text will not satisfy all four.

For example, consider a hypothetical record consisting of two conflicting payment dates. The system selects one date without documenting the conflict. A proper review of the originals returns to both dates and removes what appeared to be a factual finding from the summary. If the conflict is determinative, the final decision-maker must resolve it through some mechanism provided for in the case's applicable procedure. No comparative inference about either country's actual practices can be inferred from this illustration. Instead, it demonstrates a falsifiable test of the protocol: an independent reviewer can verify whether or not the conflict remained apparent, who resolved it and whether the published rationale is based upon the judicial evaluation.

However, the disclosure difference alters the recommended institutional policy for disclosure in either jurisdiction. In Colombia, an institutional authority may not adopt an institutional materiality standard that supplants express standards established in the

agreement for purposes of expediency. In Brazil, the distinction between decision text and logs precludes treating silence in a decision as evidence that no record is required. This book recommends additional scrutiny of a functional form in either context: when an AI contributes material propositions that were not considered by the parties during their deliberations, then courts should examine whether there exist provisions within applicable procedural entitlements to respond thereto. There is no assertion here that both AI instruments require equivalent party participation procedures.

Furthermore, the difference illustrates limitations in evaluating transparency based solely on proportions of decisions referencing AI. For instance, a ratio could report compliance with one system's drafting requirements while failing to report compliance with another system's recording protocols. A more adequate comparative record involves distinguishing among institutional publicizing practices, case-by-case disclosures to parties and internal traces of usage; none of these provide equivalencies for each other and thus enable rational explanations of what each legal design renders visible rather than equating visibility with each other form of protection.

Similarly, visibility should not serve as a panacea for unauthorized functions by AI systems. Even detailed documentation regarding how a machine-generated credibility conclusion came about may transgress boundaries of acceptable assistance. Accordingly, the policy categorizations set forth in Table 2.3 operate sequentially: authorize use of a particular function; establish controls commensurate with that function; and thereafter determine whether recording and disclosure mechanisms are sufficient to limit potential misuse of those functions. Thus, this sequential ordering represents the primary outcome of this comparison. Furthermore, this ordering transcends these two jurisdictions; as similar terminology regarding accountable AI may obscure fundamentally disparate allocations of decisional and informational accountability.

2.4 Legal Aid, Language Access, and Public Legal Information

Information is one stage in access to a remedy

A service can produce more legal information without improving a person’s ability to obtain a remedy. Information must be understandable, relevant to the correct jurisdiction and problem, and connected to a feasible next step. The person may still need representation, interpretation, documents, a recognised submission, or urgent consideration by a competent authority. Evaluation should therefore follow the user’s legal objective beyond the answer displayed on a screen. A high volume of completed conversations is not equivalent to a high number of problems meaningfully addressed.

The service framework treats an answer as one step in a legal process. Its normative justification is that a service inviting reliance should retain a practicable way to deal with foreseeable consequential error. The required response depends on the service's

actual function: explaining a court's location, preparing a pleading and screening eligibility for representation require different safeguards. The framework is assessed below through a constructed workflow and explicit release conditions, rather than through an assumed link between conversational fluency and access to justice.

The framework begins by separating three outcomes. Information access concerns whether a person obtains a relevant explanation. Procedural access concerns whether the person can take the necessary step, such as contacting an appropriate provider or presenting a matter. Effective access concerns whether the process can consider the person's position and provide the remedy available under the applicable law. A service may improve the first without improving the second or third. The evaluation should identify the level it actually measures rather than treat the most ambitious outcome as established.

General information and individualised advice

General information explains a procedure, institution, or legal topic without necessarily assessing what a particular person should do. Individualised advice connects law and facts to the person's circumstances, choices, and consequences. The boundary matters, but its professional and regulatory meaning must be established in the relevant jurisdiction. An interface cannot settle the question simply by calling every answer "information." Its actual behaviour, the expectations it creates, and the legal character of the service require examination.

A system that states where a court is located performs a different function from one that selects a cause of action, advises that a deadline has expired, or recommends abandoning a claim. The latter outputs can materially influence conduct even when accompanied by a disclaimer. The service protocol should state which functions are permitted and which require a qualified person. A request outside scope should lead to an intelligible explanation and a feasible referral route, not a confident answer generated to avoid leaving the user dissatisfied.

The analytical question is what the service does when a user's problem exceeds those limits. A disclaimer should be supported by operational behaviour: the system should refrain from presenting an unverified deadline or personalised conclusion as dependable advice, identify the relevant official source where available, and direct consequential uncertainty to a competent person. Such design does not replace the applicable law. It prevents the service's own description from being contradicted by the reliance its output invites.

Document preparation and deadline control

Document preparation can help a person organise an account, identify missing information, or complete an approved form. It can also change the substance of that account. The user should be able to review the original narrative alongside the prepared text and correct material changes. A system should not invent dates, identify a defendant without confirmation, add an admission, or select relief merely because the form requires a completed field. Missing information should remain visibly missing until it can be supplied or appropriately addressed.

The protocol should distinguish a draft from a lodged document and explain what remains to be done. A prepared form is not proof of valid filing, service, or compliance with a court direction. Where a service submits material, it needs a separate authorised process for checking the operative version, confirming the recipient, recording receipt, and communicating status. Where it does not submit, that limitation should be clear at the point the person receives the document, rather than hidden in a general notice displayed at the start of a conversation.

Deadline handling requires more than generating a calendar date. The relevant starting event, method of service, applicable rule, exceptions, and procedural context may matter. The service should identify which inputs are known and which require confirmation. A hypothetical user may describe the date printed on a letter when the relevant procedure turns on another event. A fluent answer cannot resolve the missing fact. When a consequential deadline cannot be verified, the proposed protocol requires prompt referral and a clear statement that the service has not established the time available.

The consequence of error should determine the safeguards. A wrong office opening time can usually be corrected differently from an answer that wrongly states no remedy remains available. The service should test the latter class of output directly and avoid treating it as an ordinary informational response. A user should not be discouraged from obtaining qualified assistance because the system has assigned a low likelihood of success. Initial information and screening must remain distinct from a legally or professionally authorised decision about a person's case.

Language access requires meaning to survive the whole process

Language accessibility means being able to receive and act upon a response, not merely selecting a language option. The service has to analyze the language of input, the account established based on the user's input, how options were explained to the user, the language used in the generated document, and how the user will go about contesting an error. The user might be able to ask a question in one language, yet be unable to check a document generated in another. In other words there is no automatic provision for a user to have their document reviewed by someone who can understand the content of the document when the translator provided in this instance did not use a defined method of reviewing the users' input.

Section 35(3)(k) of South Africa's Constitution protects an accused person's right to a trial in a language the person understands or, where that is not practicable, to interpretation. Section 35(4) requires information that must be given under section 35 - not only subsection (3) - to be communicated in a language the person understands. Section 35(3)(g) provides for state-assigned legal representation at state expense if substantial injustice would otherwise result, together with prompt information about that right (Constitution of the Republic of South Africa, 1996). These provisions do not establish that an automated translation tool or chatbot alone fulfils the relevant right.

The proposed language assessment approach begins with understanding the consequences of the language. Therefore, the proposed approach is to assess how long it takes to complete a task in each language, how conditional language is structured, how discretionary vs. mandatory language is expressed, how institutions are referred to by name, what remedies are available to the user, and what distinguishes asking for relief from actually obtaining it. In addition, a qualified reviewer must compare the original version of the document with its translated version(s) to ensure that both versions meet the task that the user is trying to accomplish. Back-translation may expose discrepancies, but apparent agreement is not a substitute for competent comparison with the original.

In assessing the forms of language that the service is designed to accommodate, consideration should be given to spelling variations, mixed languages, spoken responses where applicable, and specialized terminology that would likely be unknown to non-experts. These are some examples of proposed testing criteria. Claims that a particular model fails these types of tests are not part of the proposed testing criteria. If evidence does not demonstrate fitness for a particular language or communication channel, then the service must state that limitation and offer an alternative. A communication channel cannot be represented as accessible simply because the system is capable of producing grammatically correct text in that channel.

Consider a hypothetical mistranslation of a request to delay enforcement. The translated answer refers to cancellation of a remedy instead of temporarily suspending it. The user may pursue incorrect relief or misunderstand what an affirmative court order

would achieve. Identifying the erroneous word in the interface is only a first step. The service must also identify which documents or communications are affected by the prior wording, determine whether a party took action based on that wording, and implement an adequate corrective measure. As a result, language quality must play a role in identifying necessary corrections in the remedy process, in addition to technical evaluation.

Screening, urgency, and referral

Legal-aid screening must not passively impose unauthorised eligibility criteria on the applicant. A missing document, unknown address, or incomplete personal history (narrative) are distinguishable from a well-established basis of ineligibility. When a screening decision impacts accessibility, the individual should have opportunity to explain additional circumstances beyond those defined within the standard categories and to seek a review by the decision-maker with the authority to apply the applicable standards.

It would be better to determine urgency based upon the potential consequences of delaying intervention, rather than solely through a restricted vocabulary. The proposed escalation categories include a possible impending procedural deadline; current detention of liberty; a threat of loss of home or vital services; immediate risk to personal safety as a result of a legal issue; and situations which involve a child or other persons who lack meaningful ability to participate in the legal process. These categories represent conservatively designed service entry points, not statements that all such issues meet equal levels of legislative priority or entitlement under the laws of every jurisdiction.

When an urgency condition is triggered, the goal of the system should be to get human attention involved in the matter. As much as possible, the system should only ask sufficient minimal questions to enable a safe and effective transfer to an appropriately qualified provider; indicate clearly what type of assistance is requested; and avoid providing assurances regarding the outcome. For example, someone reporting danger may also need an emergency response/ protection service; the legal information pathway should not impede this path while seeking to complete an extensive survey / interview.

Establishment of local partnerships, operational hours, legal responsibility, and emergency plans are necessary prior to establishing a promise of escalation. A referral represents more than a connection. The proposed protocol indicates whether the referred-to provider meets suitability for addressing the issue; whether contact occurred; whether the issue was accepted or directed elsewhere; and what action(s) the user should take. The applicant's personal information should only be transmitted via an approved arrangement with adequate protective measures relative to the level of vulnerability present.

If the referred-to provider has no capability to assist, the original service should not record that the case has been addressed simply because a notice has been sent. Rather, the originating service should have an alternate method for referring the applicant or an honest acknowledgment of what services are unavailable.

The strongest argument against developing such systems is based on resource limitations: Legal aid organisations may not have the resources to immediately provide professional help for all requests for clarification. Prior to launching the service, these limitations will influence scope, number and qualifications of personnel, operational hours and priorities. Thus, while there are some types of limited services that could still be helpful if they offer reliable transitions to qualified providers, advertising a broad-based service that has no means to respond to obvious urgent cases will likely engender unrealistic expectations that cannot be fulfilled. In evaluating such systems, comparisons should be made to alternative approaches with available resources (e.g., a physical intake desk or pre-approved print materials), rather than assuming that either full-scale automated systems or no service at all are viable options.

Providing Access - Assisted and Non-Digital

Access provided through assisted mechanisms should permit individuals to utilise the service without relinquishing unnecessary control over their account. A helper might read aloud what is presented visually (text); describe an interface; or facilitate participation during calls, while enabling individuals to verify information submitted. Privacy and safe communication protocols require special consideration when individuals share devices or contact information. To ensure privacy and safe communication protocols are implemented when using devices or contact information are shared, services must identify how users can express concern for safe channels and prevent sensitive information from being transmitted through routes controlled by others. These procedures should be tested with anticipated users wherever feasible.

Documentation Requirements

Documentation requirements should align with stages of processing. Documentation required for general information typically does not equate to documentation requirements for applications requesting funding for representation in courts or filings related to litigation. The service should not collect a comprehensive identification portfolio simply because the standard platform allows it. When a subsequent stage legally necessitates specific documentation, applicants must be informed of the necessity for such documentation; the availability of reasonable alternatives; and who can entertain exceptions. Missing credentials should not automatically convert into judgments that render an applicant's legal problem unfounded.

In assessing non-digital access routes, only what they make accessible counts. A phone line; a local desk; printed forms; or intermediary arrangements are only valuable if they connect directly to a substantive service capable of handling corrections. Services

that claim to provide an independent route by merely directing individuals back to the website do not provide true access. Organisations must keep track of time; cost; level of assistance offered; and procedural choices available through each access route. Organisations must refrain from making unsubstantiated claims that all routes already provide equitable access.

Evaluating service outcomes and applying the protocol

Evaluation should distinguish a response delivered, a response understood, a document prepared, a submission accepted, a referral received, and a problem considered by an appropriate authority. These are separate events. The institution can assess the stage it controls while documenting dependencies beyond its control. It should not claim that a legal remedy was obtained solely because a user received an answer. Nor should lack of a favourable outcome automatically imply service failure where the process accurately explained the law and enabled a fair consideration of the matter.

Table 2.4 is the proposed service protocol. Before use, the provider should identify its legal scope, competent professional supervisor, available partners, languages and channels actually tested, and the arrangements for urgent cases. The record should include limits as well as capabilities. Changes in a relevant rule, service partner, form, or contact route require an update to the user-facing service and the staff instructions that support it.

Table 2.4. Public legal-information and legal-aid service protocol

Service stage	Minimum proposed control	Human route and evidence of completion
Scope and entry	State the jurisdiction, purpose, permitted information functions, advice limits, languages, channels, and operating arrangements.	Professional supervisor approves scope. Record tested functions and a feasible route for out-of-scope requests.
Safe and assisted access	Ask about a usable and safe communication route; collect only information justified for the stage; offer supported access.	Access lead arranges an alternative without exposing unnecessary sensitive detail; record user preferences appropriately.
Information and sources	Use identifiable material; state relevant limits; avoid presenting unverified personalised conclusions as dependable advice.	Qualified reviewer addresses consequential uncertainty. Retain source, version, and date of content review.
Documents	Preserve the user's account; verify identities and material facts; distinguish draft preparation from submission and service.	User and responsible reviewer approve substantive changes. Record the final version and what remains to be done.

Table 2.4. Public legal-information and legal-aid service protocol

Service stage	Minimum proposed control	Human route and evidence of completion
Deadlines and procedural status	Establish the rule, starting event, known facts, and any necessary qualification before giving a consequential date.	Prompt human referral if unresolved; retain the basis rather than a generated date alone.
Language	Test consequential legal meaning, not fluency alone; retain originals and identify provisional or approved translations.	Appropriate linguistic and legal review resolves material uncertainty; document test scope and known limitations.
Eligibility screening	Apply only approved criteria; distinguish missing information from established ineligibility; allow contextual explanation.	Competent officer decides consequential exceptions or refusal under the governing framework; preserve reasons and review route.
Urgency	Identify potential loss of effective relief, imminent deadlines, detention, immediate safety issues, and other locally defined triggers.	Route promptly to the appropriate human or emergency channel. Record acceptance, next step, and any unavailable assistance.
Referral	Confirm suitable provider, lawful information transfer, operating arrangements, and whether the receiving service accepts the matter.	Referral owner follows failed contact or refusal through an available alternative; a sent link is not recorded as resolution.
Correction and evaluation	Identify affected users and documents; distinguish information received, action completed, and access to consideration or remedy.	Service owner coordinates notice and repair, records unresolved consequences, and updates recurring failure controls.

Urgency triggers and alternative routes require identified capacity and partner arrangements before launch; §2.5 tests the handoff and correction sequence.

The central test is whether a person can move from an intelligible account of the problem to an appropriate and usable route for action. More information can be valuable, but it should be described as information unless the evidence establishes more. A legal-aid service improves access to an effective remedy only to the extent that its contribution survives the subsequent steps of communication, preparation, referral, participation, and correction.

2.5 Professional Duties and Changes in Legal Work

Responsibility should follow the task and the power to act

An AI-assisted team needs an allocation of responsibility that describes people's work rather than disperses it across a collective label. Naming AI as responsible does not identify who selects the task, checks the output, authorises release or corrects an error. The professional codes examined below attach duties to practitioners. The chapter translates those duties into operational responsibilities while preserving the difference between a proposed internal arrangement and the legal attribution of liability.

Operational accountability and legal liability are not identical. A team may allocate review to a particular lawyer, while the applicable law separately governs the responsibility of the practitioner, employer, institution, supplier, or other actor. The presence of human review does not establish that only the reviewer can ever be liable; the use of a supplier does not establish that the supplier bears every consequence. Liability depends on the relevant duty, conduct, causation, and remedial framework. The policy below names operational owners without pre-judging those jurisdiction-specific questions.

Professional sources must be identified by role. Lawyers, clerks, paralegals, translators and court employees are not interchangeable legal actors. The comparison in this section concerns practitioners governed by the South African Legal Practice Council Code of Conduct and Brazil's OAB Code of Ethics and Discipline. It examines a shared task - using an external system to prepare confidential client work - without assuming that either professional code regulates every court employee or supplies the client's complete civil remedy.

Confidential drafting under two professional codes

South Africa's Code of Conduct links client interests to duties to the court, justice and law; it protects privilege and confidentiality, preserves independence, requires competent and timely work, and addresses unnecessary work and reasonable fees (Legal Practice Council, 2019, paras. 3.1, 3.3, 3.6, 3.9-3.13). These provisions concern the practitioner's conduct. They do not turn a supplier's assurance of accuracy into professional verification, nor make a client's desire for speed sufficient justification for releasing work the practitioner cannot competently assess.

Brazil's OAB Code similarly protects independence and addresses truthful professional conduct, client information, confidentiality and fee arrangements. Its confidentiality provisions apply without a client's request for secrecy and recognise exceptional just cause; its fee provisions require clarity about the engagement and identify contextual factors relevant to moderation (OAB Code, 2015, arts. 2, 4, 6, 9, 35-38, 48-49). The

public-order character of confidentiality makes it particularly inappropriate to treat a generic acceptance of software terms as a complete analysis of the lawyer's duty.

The matched problem is a draft pleading prepared from a client's confidential narrative using an external service that may retain inputs. In both settings, the practitioner must first characterise the disclosure and determine whether the chosen arrangement is compatible with the applicable duty. This conclusion does not depend on whether the generated pleading is correct. Confidentiality can be compromised at input, before any output exists. An output-only review therefore leaves an important part of the professional task unexamined.

A second problem arises if the system proposes a plausible allegation not contained in the instructions. The practitioner cannot resolve it by checking the cited law: the deficiency concerns factual authority to make the allegation. The workflow should preserve the client account, flag the addition and obtain clarification before adoption. The inference drawn here is operational, rather than a claim that either code expressly mandates a particular software ledger. A record of instructions and corrections is evidence that the professional exercised judgment; it is not a substitute for doing so.

The comparison also sharpens the fee question. The codes supply professional constraints, while the engagement supplies the particular charging basis. Neither code justifies inventing hours because an earlier manual process would have taken longer. Equally, a fixed-fee arrangement cannot be analysed by assuming that every saving belongs automatically to one party. The relevant questions concern what was agreed, whether required information was supplied, whether the fee is permissible, and whether experimentation or avoidable rework has been passed on without a proper basis. The policy should record those questions separately from the tool's subscription price.

The shared result is that adoption changes the evidence needed to demonstrate existing duties, even without an AI-specific rule. The difference between the codes should remain visible: their wording, scope and disciplinary arrangements are not interchangeable. The comparison supports task-specific records and training, not a universal code assembled by selecting the broadest phrase from each jurisdiction. This is also why the operational owners in Table 2.5 do not predetermine civil liability or eliminate duties owed by a firm or supplier.

A named chain of review

At the beginning of a matter, the responsible professional should approve the task and the permitted use of the output. This involves more than approving a product. The same system may be acceptable for rearranging public material and unsuitable for processing a confidential file or producing an unreviewed advice letter. The task record should identify the intended recipient, the consequence of error, the authorised inputs, and the required reviewer. A change in purpose requires a new assessment even when the software remains unchanged.

The person operating the tool should record enough information for another reviewer to understand what was requested and produced. The legal reviewer checks authority, interpretation, procedure, and the relationship to instructions. A language reviewer checks consequential translation where the task requires it. A records custodian preserves the approved version and the relevant evidence of review. The final professional or institutional owner authorises communication or filing. These roles can be combined in a small team, but the combination should remain explicit and supported by the necessary competence.

Review responsibilities should be bounded as well as assigned. A technical specialist can explain system operation without being qualified to settle a legal proposition. A bilingual colleague can notice an obvious translation defect without being qualified to certify a legally consequential rendition. A junior lawyer can perform a source check without having authority to approve the final advice. The policy should identify the matters each reviewer can resolve and those that require escalation. Otherwise a chain of signatures may conceal that nobody accepted responsibility for the central issue.

Supervision requires access to the work actually performed. A senior professional who sees only a polished final document cannot readily determine whether a source was opened, whether confidential material was entered, or whether an uncertainty was removed without justification. Review records should therefore be proportionate to the task: a proposition ledger for research, a redline and instruction record for contracts, a sampling and exclusion record for document review, and an appropriate source-and-correction record for translation. No single generic attestation can demonstrate all these activities.

Competence must be demonstrated in the relevant task

Technical competence should be defined in terms of what the professional needs to recognise and control. It does not require every lawyer to design a model. It does require the user to distinguish generated text from a retrieved source, understand the stated scope and limits of the tool, recognise relevant input restrictions, and know when to stop. A declaration of familiarity with AI or possession of a training certificate cannot alone demonstrate those abilities. The assessment should examine work performed under conditions comparable to practice.

A proposed competence exercise could give the practitioner a draft containing a genuine but misapplied authority, an unsupported factual assertion, an omitted exception, and a confidentiality problem in the proposed input. The practitioner would identify the defects, retrieve and interpret the sources, decide which material may be processed, and produce a corrected version with an explanation of remaining uncertainty. This is a proposed assessment design, not an administered test or evidence of actual competence levels. Its advantage is that it evaluates judgment rather than confidence in operating an interface.

Competence should be task-specific. Approval to use an internal summarisation tool does not establish competence in litigation drafting, multilingual advice, or large-scale disclosure review. A translator needs the ability to preserve legally consequential meaning and identify when legal clarification is required. A clerk needs to distinguish a draft recommendation from an operative direction. A supervisor needs to assess whether the checking process itself was adequate. Different tasks can require different evidence and different levels of independent review.

Continuous learning should respond to change rather than become a recurring general lecture. A new model, retrieval collection, data setting, intended use, procedural rule, or incident may create a need for renewed assessment. Staff should know how changes are communicated and when previous approval no longer covers the task. The organisation should retain evidence of the updated instruction, the practical exercise where appropriate, and the person authorising resumed use. The purpose is not to maximise training hours but to maintain demonstrable ability to perform the approved work.

Confidentiality, candour, and client communication

Confidentiality responsibilities extend through the team's full processing chain. Approval should address the institution's accounts, permitted data classes, retention settings, external access, and any outsourcing arrangement. Individual staff should not bypass those arrangements through personal accounts because an approved system is inconvenient. The policy should provide a workable way to seek help or use a non-AI method when the approved environment does not meet the task. A restriction that makes necessary work impossible may encourage undisclosed use unless the institution provides an alternative.

Candour concerns the accuracy of the professional's own representations. A source should not be described as checked when nobody examined it, and a generated passage should not be defended solely because it appeared in a reputable product. The particular obligations owed to a court, opponent, regulator, or client must be identified from the applicable rules. As an organisational control, the policy should prevent unverified work from acquiring a false review status and should require material uncertainty to reach the person responsible for the final communication.

Client communication should focus on matters material to the engagement rather than promise disclosure of every software operation regardless of significance. The responsible professional should consider whether the proposed use changes the handling of confidential material, the scope of work, the identity of those reviewing it, cost arrangements, or the limitations of the advice. Any consent or disclosure requirement imposed by applicable law or the engagement takes priority. Where consent is sought, the explanation should describe the actual activity and its relevant consequences, not rely on an abstract question about whether the client accepts "AI."

Disclosure to a client and disclosure to a court are separate decisions. A client's agreement to use a tool does not establish compliance with a court direction, and a court-facing statement does not necessarily inform the client adequately about processing or cost. The policy should name the person who assesses each duty and records the basis of the decision. Where an incident changes the accuracy, confidentiality, or usability of delivered work, the communication assessment should be reopened rather than treated as settled by the initial engagement.

Billing and the accounting of assistance

AI-assisted work can change the relationship between drafting time, checking time, technology cost, and professional value. Billing should reflect the agreed basis and applicable professional rules rather than an assumed entitlement to charge the time a manual method might have taken. Under an actual-time arrangement, the proposed policy records the time actually spent on authorised work, including necessary verification, and does not enter hypothetical hours as though they were worked. This is a model accounting safeguard, not a statement that identical billing rules apply across all six jurisdictions.

Fixed or otherwise agreed fees raise a different question. A reduction in production time does not automatically determine the fee where the engagement lawfully adopts another basis. The professional should examine the agreement, required communication, permissible expenses, and any rule governing fairness or reasonableness. Technology charges should not be concealed, duplicated, or attributed to a client without an appropriate basis. The chapter does not prescribe a universal fee reduction or a universal right to retain efficiency gains.

Verification time should be visible enough to distinguish necessary professional work from avoidable rework caused by an unsuitable tool. The organisation may need to bear some training or experimentation cost rather than silently pass it to a client; the appropriate allocation depends on the governing rules and engagement. Recording that distinction also improves adoption decisions. A tool that repeatedly produces unusable drafts may increase billed activity without improving service. That is an operational finding to investigate, not a reason to equate more recorded time with greater professional value.

Outsourced work and supplier dependence

Outsourcing does not eliminate the need to know who performs consequential work. The organisation should identify the service provider, permitted inputs, access arrangements, subcontracting, review responsibilities, and method for obtaining relevant records after a problem. It should distinguish the supplier's technical assurances from the professional's acceptance of a particular output. A contractual allocation of responsibility can support correction and recovery, but does not itself determine the rights of a client or other affected person who is not a party to the contract.

The policy should cover supplier changes that affect practice. A revised model, a different data-processing arrangement, loss of source access, or the removal of a useful control may invalidate an earlier assessment. Staff need an approved route for pausing work, retrieving records, and using a substitute process. Essential client or court material should not remain accessible only through a supplier account that may later become unavailable. Continuity is part of supervision because a professional cannot review or correct work that the organisation cannot recover.

Early-career training and work allocation

Task assistance does not by itself establish a forecast of employment displacement. A workforce claim would require evidence about demand, task substitution, new work, pricing and organisational responses. The question examined here is narrower and normative: how should a team distribute research, drafting, checking, client contact and supervision so that it can fulfil its existing professional responsibilities? That inquiry can guide present practice without predicting the number of future legal jobs.

Early-career work serves both production and learning. A practitioner who only checks the final output may encounter a different learning problem from one who must locate and compare authorities before drafting. The concern is a causal hypothesis, not a measured loss of competence: if independent source selection disappears from training, the organisation needs another way to demonstrate that its future reviewers can perform it. A practical exercise that requires explaining a rejected authority tests this capacity more directly than a count of documents completed.

A proposed training policy can preserve selected independent research and drafting exercises, require explanation of material corrections, and include supervised client-facing work where appropriate. It can rotate staff through preparation, verification, and incident review rather than assigning all checking to the least experienced person. Supervisors should assess reasoning and source use, not only production volume. The objective is not to retain every old task unchanged, but to ensure that faster production does not conceal a failure to develop the capacity needed to supervise it.

Work allocation should also recognise the pressure created by increased volume. A team may save time preparing each draft yet receive many more drafts to verify. The relevant capacity measure is therefore the amount of consequential work reviewers

can properly examine, not the amount the system can generate. Institutions should document queue length, review time, correction patterns, and the tasks returned for further work. Those records can identify whether a bottleneck has moved rather than disappeared. They do not, without further analysis, establish the broader employment effect.

Escalation, incidents, and a professional-use policy

Staff should have a route to raise uncertainty without needing to prove misconduct or a technical failure. Escalation triggers include an untraceable source, inconsistent facts, an apparent confidentiality issue, a consequential translation problem, an output outside approved scope, and insufficient time for the required review. A supervisor should decide whether the task can be corrected, narrowed, reassigned, or performed without AI. The policy should protect the integrity of the review process rather than reward silent acceptance of questionable work.

When a material error is discovered after release, the organisation should preserve the relevant version and record, identify affected recipients and matters, assess immediate consequences, and use the appropriate correction procedure. The review should distinguish individual oversight from recurring weaknesses in source access, configuration, training, or workload. A correction in one file should inform a search for similarly affected work where justified. Deleting an embarrassing output without assessing reliance would leave the institution unable to establish what needs repair.

Table 2.5 consolidates the professional-use policy. The responsible organisation should add the governing professional instruments, engagement terms, court directions, named staff, and review dates before implementation. Its allocations concern operational conduct; they neither exclude supplier responsibility nor settle professional, civil, disciplinary, or public-law liability in advance.

Table 2.5. Professional-use policy and allocation of review responsibilities

Policy responsibility	Named operational owner	Demonstration, record, and escalation
Approve task and tool	Responsible professional or institutional matter owner.	Record intended use, inputs, recipient, consequence, approved environment, governing rules, and non-AI alternative. Reassess changes in scope.
Operate within scope	Authorised practitioner or staff user.	Retain task instructions, output, relevant version, and uncertainties. Stop and refer activity outside approved use or input permissions.

Table 2.5. Professional-use policy and allocation of review responsibilities

Policy responsibility	Named operational owner	Demonstration, record, and escalation
Verify legal work	Reviewer competent in the relevant field and jurisdiction.	Complete source and interpretation checks, contrary-material review, and procedure assessment. Do not attest to checks not performed.
Review language	Appropriately qualified language reviewer, with legal support where needed.	Demonstrate preservation of consequential meaning; record unresolved terms, source text, corrections, and approval status.
Protect information	Matter owner with designated information-governance support.	Approve access and processing conditions, accounts, retention, outsourcing, and permitted data. Refer uncertainty before upload.
Authorise release	Professional or institutional actor responsible for advice, filing, or communication.	Approve the exact version, purpose, recipient, and limitations; confirm that required reviews are complete.
Assess disclosure and consent	Responsible professional, with procedural or ethics advice when necessary.	Separate client and court obligations; identify governing source, material consequence, explanation, and decision. Reopen after a relevant incident.
Record fees and costs	Engagement owner and billing supervisor.	Apply the lawful agreement and rules; record actual time where relevant, authorised expenses, and allocation of training or rework.
Demonstrate competence	Practice supervisor or training lead.	Use task-based exercises and observed work; record approval limits and reassessment after material changes. A certificate alone is insufficient evidence.
Supervise outsourcing and continuity	Supplier/contract owner with the matter supervisor.	Identify providers, access, evidence rights, changes, record export, and alternative operation. Escalate lost access or unapproved material change.

Table 2.5. Professional-use policy and allocation of review responsibilities

Policy responsibility	Named operational owner	Demonstration, record, and escalation
Preserve records and correct errors	Records custodian and responsible professional.	Protect relevant versions; identify affected recipients and matters; coordinate the appropriate correction and notification process.
Review work allocation and learning	Team leader or institutional management.	Examine review capacity, correction patterns, learning opportunities, and workload. Do not infer staffing effects from generated volume alone.

Operational ownership does not settle legal liability. Combine roles only where competence and capacity are demonstrable; add actual names, applicable instruments, substitutes, and review dates.

Professional competence in an AI-assisted team is demonstrated when the team can show what it checked, why it accepted the result, how it communicated relevant limits, and what it did when the work was wrong. Responsibility should be assigned to people able to perform those acts and escalate what they cannot resolve. A policy that simply declares human accountability, without the time, records, skills, and authority needed to exercise it, is an incomplete organisational arrangement.

A completed protocol exercise for a small legal-information service

The following constructed exercise tests the internal coherence of Tables 2.1-2.5. It is deliberately a proposed service, not a report of a deployment. A small legal-aid office has one supervising lawyer, one intake worker and access to a contracted interpreter during advertised hours. It wishes to use an approved system to turn a user's narrative into a chronology for the lawyer. It will not determine eligibility, calculate an unverified deadline, file a document or provide an autonomous merits assessment. Restricting the task is the first resource decision.

The completed task entry is: purpose, preparation of a provisional chronology; recipient, supervising lawyer; operator, intake worker; release authority, supervising lawyer; permitted input, the minimum account accepted through the approved confidential channel; fallback, direct intake using the same structured form. The case identifier is L01, the task identifier T01, the source account S01 and the proposed chronology C01. These identifiers are examples of the common links needed across the records, not fictitious dates or evidence of completed approval. An office may use a paper index if it preserves the same relationships.

The intake worker receives two uncertain dates and a narrative suggesting that an event may occur before ordinary review is available. She records the reported event, the user's words and the uncertainty without inventing a legal deadline. She routes the matter to the supervising lawyer for urgency assessment; if the lawyer is unavailable, she uses an established partner or emergency route appropriate to the problem. If no competent route is available, the office must state that limitation. An unstaffed escalation button cannot support a promise of urgent assistance.

The system generates a chronology that changes "I was informed that my payments may cease" into "my payments ceased". The intake worker flags the discrepancy against source account S01 without altering that account. Chronology C01 remains provisional. The lawyer checks S01, removes the unsupported assertion and approves corrected version C02. Unresolved legal or linguistic questions go to a competent reviewer. The release condition is that C02 faithfully represents the account, preserves uncertainty and is suitable for the lawyer's stated next step.

A minimum amount of documentation to provide for this function includes authorized scope, source account, operative timeline, material modification, reviewer and timing of review, undecided issue and next accountable party. There is no requirement to include every unaffected paragraph or retain irrelevant personally identifiable information. Additional documentation occurs whenever enhanced documentation is required due to the output providing legal interpretation to an outside entity receiving protected information, a disputed translation occurring or prior communication regarding an error occurred. Each additional category of documentation provides a means whereby a reviewer can address a given query. Documentation without a purpose may introduce additional avoidable confidentiality obligations and retention requirements.

For example, assume the erroneous version was transmitted to a partner. Correction involves identifying that partner, transmitting C02 and determining if any subsequent actions were predicated on C01. The partner may have to reassess the appropriateness of its own actions under its jurisdiction; correcting the originating office's file will not accomplish this automatically. Therefore, the closure entry differentiates among correction sent, correction acknowledged and consequential action completed. Although a case may remain undecided after completion of stages one and two, the protocols track transitions between offices.

The exercise has three stop conditions: input permission is unresolved; a material assertion cannot be checked in the available time; or the service cannot provide the escalation it has promised. The office should pause the affected use, adopt an available lawful alternative, or narrow its service commitment. These decisions turn on the consequence of reliance and the capacity to correct it, rather than on an undifferentiated checklist score.

Practitioners should evaluate feasibility before expanding the scope of tasks. Offices document employee time spent performing direct intake and assisted intake functions

including review, language support, unsuccessful referrals and correction. They then determine whether the system decreases workload at acceptable levels of quality or simply redistributes efforts from drafting to reviewing. Users/staff should test whether the status indicators and next steps are understandable. While this tier does not demonstrate an empirical savings, it demonstrates a repeatable testing plan and clearly articulates which type of evidence is required to assert such savings.

Additionally, this lowest-tier of implementation describes how roles may be appropriately aggregated. For example, one attorney can serve as both Source Reviewer and Release Owner for a limited scope case; the record must indicate that as an aggregate role, not as separate reviews. However, an individual cannot utilize a second signature to represent missing linguistic competency. When there are insufficient resources / expertise or time, a reasonable response is to implement a narrowed scope task or reliable referral or alternative processes. Proportional application of administrative burdens supports efficiency while maintaining control over those factors upon which reliability of a specific service relies.

Chapter 2 Key Arguments

The chapter establishes that assistance is a relationship among an output, a legally assigned task and an accountable user. Verification must address source, interpretation and intended use; administration must be followed into its procedural consequences; and professional responsibility must be supported by access, competence and time. These propositions explain why neither a genuine citation nor a human signature is a sufficient release condition.

The Colombia-Brazil comparison shows that a shared commitment to human responsibility can coexist with different disclosure architectures. The professional-code comparison shows a complementary continuity: existing obligations concerning confidentiality, independence and competent work remain material when the means of production change. The appropriate response is neither to assume legal novelty everywhere nor to conclude that familiar duties require no new operational arrangements.

The constructed legal-information exercise demonstrates how the protocols connect in one bounded workflow, including a material correction and a failed handoff. It identifies minimum records and stop conditions without claiming that the proposed service has been evaluated. The remaining empirical question is whether an institution can meet those conditions at acceptable cost and quality. Chapter 3 develops the data and supplier records needed to answer that question and to preserve the possibility of correction.

CHAPTER

03

Data Governance and Automated Decisions

IN THIS CHAPTER

3.1 Legal Bases for Data Collection and Reuse

3.2 Legal Data, Language Coverage, and Intellectual Property

3.3 Public Benefits and Identity: Historical Administrative Analogues for AI

3.4 Automated Decisions in Employment, Credit, and Digital Platforms

3.5 Cross-Border Processing and Technology Supply Arrangements

Introduction

A legal decision can become difficult to challenge long before an algorithm produces its output. The difficulty may begin with a field that cannot accommodate a person's circumstances, a record collected for a different purpose, an unexplained match between databases, or a contract that leaves essential evidence with a supplier. By the time an applicant receives a refusal or a worker loses access to an account, the decisive information may have passed through several institutions. Each may hold part of the record, while none can explain the complete route from collection to consequence. Data governance therefore belongs within the analysis of legal authority and remedies, not outside it as a preliminary technical exercise.

The preceding chapters established two distinctions on which this chapter builds. Adopting software is not, by itself, legal transformation; transformation concerns material changes in the operation of authority, interpretation, responsibility, participation, and remedy. Equally, an output described as assistance can become practically determinative when institutional arrangements make it difficult to question. The present chapter follows the information through those arrangements. Its central question is how a record, classification, or inference becomes a reason for action, and what must remain available when its use is disputed.

South Africa's Protection of Personal Information Act 4 of 2013, commonly called POPIA, supplies the principal statutory starting point. Its architecture connects processing conditions, data-subject participation, certain automated decisions, and transfers outside the Republic. These are related but distinct subjects. A permissible transfer does not establish that the receiving institution may make a particular decision; an accurate record does not establish its relevance to that decision; and a correction procedure does not necessarily reopen the resulting administrative or contractual act (POPIA, 2013, ss. 3-5, 24, 71-72).

POPIA is examined as South African law, not as a regional code or a substitute for the domestic sources of Pakistan, India, Kenya, Brazil, or Colombia. The book's six-jurisdiction design remains purposive and function-based. This chapter uses selected South African and Brazilian provisions to show how questions must be separated, and two historically documented Pakistani administrative functions to examine the movement from information to consequence. It does not claim equivalent legal protection, equivalent institutional capacity, or equivalent deployment across the six countries.

The argument proceeds through five connected inquiries. Section 3.1 identifies the legal questions at each stage of the data lifecycle and distinguishes correction of information from reconsideration of a decision. Section 3.2 examines the construction of legal corpora, including authority, language, privacy, licensing, and protected knowledge. Section 3.3 compares household-based benefit targeting with biometric verification at payment, keeping statistical assessment separate from eligibility and authentication separate from entitlement. Section 3.4 examines employment, credit, and platform de-

cisions through the duties attached to the relevant relationship. Section 3.5 traces foreign hosting, subcontracting, evidence access, and continuity obligations.

The practical result is a linked documentation pack rather than a universal compliance score. A data-use register identifies what processing is authorised. A corpus assessment records what a system can and cannot know. A decision record connects the governing criterion with evidence, output, officer action, and review. A sector comparison identifies the duties that survive automation. A dependency register identifies the actors and records on which continued control depends. Their value lies in making disputed steps reconstructable; completing a form cannot create a power or cure a defective decision.

3.1 Legal Bases for Data Collection and Reuse

The processing activity as the unit of legal analysis

An institution rarely undertakes only one activity when it says that it is using data. A service office may collect an application, compare it with an existing register, send selected fields to a contractor, retain the response, and use the result in a decision. A development team may subsequently request the same records to train or test a model. These activities share information but differ in purpose, recipients, duration, and consequence. A defensible legal assessment identifies each activity rather than treating the original collection as authorisation for the entire technical lifecycle.

The first question is which law applies to the actor and activity. POPIA's application provisions concern recorded personal information processed by or for a responsible party with the specified South African connection. Its exclusions include information de-identified so that it cannot be re-identified and processing relating to the judicial functions of courts identified in the Constitution. The latter is especially important after Chapter 2: a court's involvement does not justify treating every judicial, administrative, employment, and supplier activity as legally identical (POPIA, 2013, ss. 1, 3, 6). An exclusion from one statute is not evidence that confidentiality, constitutional requirements, professional duties, or other applicable law disappear.

Legal categories also need to remain distinct from operational job titles. The office holding the authoritative source record may differ from the organisation determining the purpose and means of a new processing operation. A vendor may process information on instructions, while a separate institution decides whether the result affects eligibility. Calling all three a "data owner" conceals the differences. The chapter uses custodian to identify practical responsibility for a record, while statutory terms such as responsible party or controller retain their jurisdiction-specific meanings. A register should record both, without allowing a custody label to answer an unresolved legal question.

Purpose, necessity, and compulsory collection

A stated purpose must be specific enough to constrain what happens. “Improving services” provides little guidance about whether a department may compare household records, infer health conditions, train a general-purpose model, or share information with an employer. A useful purpose statement identifies the service, the information needed, the people affected, and the intended institutional action. It also identifies excluded uses. This is not a demand to predict every future technological change. It is a way to prevent a broad organisational ambition from replacing examination of the actual processing proposal.

POPIA separates minimality, justification, and purpose specification. Section 10 addresses whether information is adequate, relevant, and not excessive for the purpose. Section 11 provides several possible justifications, including consent and specified forms of contractual necessity, legal obligation, public duty, and legitimate interest. Section 13 addresses the specified lawful purpose of collection (POPIA, 2013, ss. 10-13). Their separation matters: identifying a ground for processing does not make every available field necessary, and establishing necessity for a task does not itself identify the institution authorised to perform it.

Consider a hypothetical application for an existing public service. The agency requests identity evidence needed to distinguish applicants, but the interface also requires access to a complete contact list for possible fraud analysis. The second request needs its own explanation. The fact that both fields appear on the same screen does not make their necessity equivalent. The assessment should ask what legal criterion each field helps establish, whether a less intrusive source would suffice, and whether refusing an additional field lawfully prevents consideration of the application. A technical team should not determine compulsory disclosure merely by making a field mandatory.

Consent should be examined with the same discipline. It is neither a universal prerequisite nor an unrestricted authorisation. Where the actual ground is a statutory obligation or public duty, a consent screen may misdescribe the person’s position. Where consent is relied upon, the institution must investigate whether it meets the governing requirements and covers the particular activity. A person’s agreement to receive a service is not analytically equivalent to agreement that their records become development material for unrelated customers. Nor does refusal of optional development use necessarily determine entitlement to the underlying service. The relevant relationship must be specified before consent is requested.

Reuse for training, testing, and model development

Repurposing can occur through a change in population, prediction target or institutional relationship, even when the activity keeps the same broad service label. Biega's computational analysis argues that a change of task alone is an insufficient account of purpose compatibility (Biega, 2023, §§5-6). The doctrinal question remains governed by the applicable processing law. The operational contribution is to make the changed features visible before the institution assesses compatibility, necessity and the recipient's role.

For legal institutions, development use should therefore be described separately from service delivery. Using a record to answer the person's application differs from using it to adjust a model that will affect future applications. Testing the model against historical records differs from incorporating those records into parameters or a searchable retrieval collection. Supplying records to an outside developer adds questions concerning the recipient's purpose and control. These distinctions do not establish that every secondary use requires a new consent or a new statute. They establish the questions that the applicable legal framework must answer.

POPIA's further-processing provision offers a concrete example. Section 15 requires examination of compatibility and identifies relevant factors and circumstances in which further processing is not incompatible. Its specified research-related treatment and treatment of public records have conditions; they should not be converted into an unrestricted permission for commercial model development (POPIA, 2013, s. 15). A reviewer must match the facts to the provision actually relied upon. The label "research" cannot establish that a production model, supplier reuse, identifiable disclosure, and indefinite retention all satisfy the same conditions.

A separate assessment is particularly important when the new processing changes what is inferred about a person. A record of a missed appointment may become a predictor of unreliability; an address may become a proxy for financial risk; a complaint may become a training label for suspected abuse of a service. Each transformation contains a choice about what the original event signifies. The assessor should identify that choice, its evidential basis, and its proposed legal relevance. A model can calculate an association without establishing that the association is an appropriate reason for institutional action.

The same reasoning applies when a development process removes names. Replacing direct identifiers does not answer whether the remaining information can identify a person through linkage, rare attributes, or retained keys. A limited test dataset should be assessed against its actual contents and access conditions rather than its label. Conversely, genuinely non-identifiable material may alter the legal analysis. The appropriate conclusion depends on the governing definition and evidence of identifiability, not a presumption that every derived dataset is either fully anonymous or permanently personal.

Sensitive information and differentiated safeguards

Sensitivity cannot be reduced to a single security classification. Legal systems identify particular categories and impose different conditions, while the practical consequences of combining information may extend beyond those categories. POPIA treats special personal information and children's information through distinct provisions, with prohibitions and authorisations that require examination rather than a generic consent statement (POPIA, 2013, ss. 26-35). The chapter does not import those categories into every jurisdiction. It asks the responsible institution to identify the applicable classification before deciding what collection, access, development, or disclosure is possible.

A proposed register should distinguish supplied information, observed events, professional assessments, and machine-generated inferences. A medical certificate, an attendance record, and a predicted health condition are not interchangeable merely because all can be stored as fields. The certificate has an identified author and purpose; the event records an occurrence under particular conditions; the inference depends on a method and assumptions. Treating them as equivalent facts can make later correction unnecessarily difficult. The record should preserve their different status and identify the actor responsible for each.

Restrictions also have to follow copies and transformations. A protected narrative may remain revealing after names are removed or when it is compressed into a summary. A translation can preserve the same confidential substance in a different language. A vector representation or other derived store should not be excluded from the inventory simply because it is less readable than the original document. The proposed control is an evidence-based inventory of the information and its uses, not a claim that every technical representation has an identical legal status.

Correction after information has influenced a decision

Correction presents at least three different problems: amendment of the source record, treatment of copies and derived outputs, and reconsideration of the consequential decision. The person may need all three. Correcting an address in one register does not necessarily remove a flag generated earlier. Removing the flag does not necessarily reverse a refused application. A later database update may therefore be technically successful while leaving the person's legal position unchanged. An effective procedure identifies the actor with authority at each stage and links their actions through a common case record.

POPIA supplies specific, qualified rights and duties rather than a general promise of correction. Section 24 provides grounds for requesting correction or deletion and addresses the responsible party's response. Where a change affects past or future decisions, subsection (3) requires notification to previous recipients if reasonably practicable. Section 14 separately addresses restriction in specified circumstances, includ-

ing contested accuracy while it is verified (POPIA, 2013, ss. 14, 24). These provisions should be read alongside the law governing the original decision; they are not a self-contained code for reopening every benefit, credit, or employment outcome.

A workable institutional procedure begins by recording what is disputed. A person may challenge the truth of a fact, the match to their identity, the currency of the information, or its relevance to the decision. A disagreement with an inference may concern its inputs, its method, or the weight assigned to it. Asking the person only to supply a corrected value will not resolve all these issues. The reviewing officer needs a way to record explanatory evidence and decide whether the contested information should be used at all.

The next step is to preserve the material needed to understand the earlier action while controlling further reliance. Audit preservation and continued operational use are not the same thing. A superseded value may need to remain in a protected history to explain an earlier refusal, but should not silently repopulate the live system. Recipients should receive an intelligible correction event identifying the affected record, date, and scope. The decision-maker should then determine whether reconsideration, restoration, or another remedy is authorised and required. These are linked actions, not an undifferentiated deletion request.

Model development creates an additional question. Where disputed records contributed to training, removing a current entry does not prove that their contribution has been removed from the model. The institution should establish what can be identified, what technical remedy is available, and what interim restriction is appropriate. Depending on the problem, measures could concern the retrieval corpus, a derived feature, a model version, or the use of outputs in affected cases. The chapter does not assert a universal technical ability or legal obligation to perform complete machine unlearning. It requires an honest account of the problem and of the remedy actually delivered.

Retention, deletion, and a usable data-use register

Retention requires classification by record and purpose. The authoritative application, a temporary development copy, the data used in a particular decision, an access log, and a backup may require different treatment. POPIA's retention provision includes both limits and specified grounds for continued retention; it also addresses the opportunity to access records used in decisions and destruction or de-identification when retention is no longer authorised (POPIA, 2013, s. 14). A schedule must therefore reconcile accountability with limitation rather than choose indiscriminate deletion or indefinite preservation.

The strongest objection to activity-level documentation is its cost. A small institution may lack specialist staff, and repeated reviews may delay a useful improvement. The answer is to reuse reliable documentation without reusing an earlier legal conclusion

beyond its scope. A standard record can cover repeated operations with the same purpose, actors, information, and controls. Material changes should reopen the relevant fields rather than require an entirely new institutional study. This permits proportionate administration while retaining evidence of why the processing was considered permissible.

Table 3.1 sets out the proposed data-use register. It is a completion template, not a description of an existing deployment. Each row requires an actual authority and named custodian. Retention entries should identify the governing rule and the event that starts or ends the period, rather than contain a guessed number of years. Affected-person procedures should identify the office able to act, the information it needs, and its relationship with the decision-maker. Unresolved entries remain unresolved; a technically available operation is not a substitute for a legal conclusion.

Table 3.1. Data-use register: activities, permissions, custody, and affected-person procedures

Activity and record	Purpose and authority	Custodian and recipients	Retention and control	Affected-person procedure
Collect for service	Identify service, necessary fields, enabling provision, processing ground, and compulsory or optional status.	Name collector, responsible party, authoritative source, and receiving office.	Record schedule, start event, access roles, notice version, and security requirements.	Explain consequences of refusal; identify assisted entry and source-correction route.
Clean, label, or link	Record purpose of transformation, authority to link, relevant categories, and any required prior authorisation.	Identify annotator, source custodian, receiving system, match method, and version.	Retain necessary provenance and change history; separate temporary working copies.	Permit challenge to a fact, label, match, or missing context; record dispute owner.
Train or fine-tune	State model-development purpose, permitted population and use, compatibility finding, and separate ground if needed.	Identify developer, dataset version, processors, model owner, and onward recipients.	Set development-copy retention, access limits, deletion events, and residual technical limitations.	Explain reuse where required; connect objection or correction to affected data and model uses.

Table 3.1. Data-use register: activities, permissions, custody, and affected-person procedures

Activity and record	Purpose and authority	Custodian and recipients	Retention and control	Affected-person procedure
Test or validate	Identify testing objective, necessity of identifiable data, legal basis, and limits on live consequences.	Name evaluator, test-set custodian, approved environment, and output recipients.	Record access controls, test duration, preserved results, and disposal of unnecessary copies.	Provide a route to question test-data use and correct errors that enter operations.
Deploy and decide	Link each input and inference to the service purpose and the law governing the decision.	Name system operator, source custodian, decision-maker, and case-record owner.	Preserve relevant input, rule/model version, output, officer action, and decision evidence.	Give applicable notice and reasons; identify data correction, decision review, and urgent route.
Share or transfer	Identify recipient purpose, disclosure authority, transfer route, and onward-use restrictions.	Record legal entities, locations, actual access, subcontractors, and destination records.	Set recipient retention, evidence-access duties, security terms, and return/deletion arrangements.	Identify disclosure information and the route for propagating a correction or restriction.
Correct and propagate	Record disputed matter, correction basis, interim restriction where available, and reconsideration authority.	Name source owner, affected recipients, model owner, and competent deciding body.	Preserve protected history while controlling live reliance; obtain correction acknowledgments.	Provide tracking, reasons, review outcome, and the action taken on consequential decisions.
Retain, archive, or delete	Identify record class, statutory / contractual rule, purpose expiry, and any lawful preservation hold.	Identify archive, operational copies, backups, derived stores, and disposal sign-off.	State retention event, review date, method, verification, and any justified exception.	Preserve access and pending challenges; explain disposal or continued retention as required.

Each entry identifies the actual instrument, person, date and record. Retention and processing permission require separate legal support.

The register should be approved by the institution responsible for the activity, with legal and records-management review. Its entries should link to the corpus assessment and decision record developed below. A change of supplier, purpose, population, data

category, training arrangement, or retention practice should trigger examination of the affected permissions. The governing question is not whether the institution already possesses the information. It is whether this actor may undertake this operation for this purpose, under conditions that permit the person to understand and challenge its consequences.

3.2 Legal Data, Language Coverage, and Intellectual Property

Access to law and the construction of a corpus

A legal corpus is a selected collection, not a complete account of law. Its boundaries are determined by what was recorded, made available, acquired, licensed, translated, indexed, and retained. A system using that corpus can retrieve only what its collection and processing arrangements make available. A fluent response cannot compensate for an omitted amendment, an unreported decision, or a form of legal knowledge that was never represented. Corpus construction consequently affects the authorities that users encounter and the kinds of claims the system is capable of recognising.

Access to legal information and permission to reuse it are separate questions. This chapter does not infer an unlimited training entitlement from the public character of law. Its narrower claim is functional: a person cannot reliably challenge a legal proposition whose authority, applicable version or relevant qualifications remain inaccessible. The corpus assessment therefore joins access with provenance and rights analysis. A collection can improve discovery while still excluding sources necessary to resolve the user's actual problem.

Access to a website is only the beginning of that process. A person may be able to read a judgment without being able to identify its procedural status, later treatment, or application to a different dispute. A developer may obtain a large collection without knowing whether it includes all relevant courts or periods. A system may answer questions in a language not represented in its underlying authorities. These are different limitations and need different responses. An access claim should specify the audience, material, function, and conditions rather than rely on the number of available documents.

The corpus assessment should therefore begin with an inventory of source classes. Legislation, judgments, procedural rules, contracts, commentary, professional forms, and community accounts do not possess the same authority or require the same permissions. Their combination in a single search interface does not remove these distinctions. The interface should make the status of the retrieved material visible, while the underlying record preserves its source, date, rights, and changes. The purpose is to prevent a convenient collection from becoming an unexplained hierarchy of legal knowledge.

Authority, currency, and the unit of verification

The unit of verification should be more precise than the document title. A statute may have different commencement dates for different provisions. A judgment may contain several opinions and distinguish findings from submissions. A consolidation may incorporate an amendment while also preserving superseded wording for reference. A procedural form may remain accessible after the rule it implements has changed. The system needs to identify the particular version and passage supporting its answer, not merely the existence of a document with a plausible name.

A workable provenance record connects each item to its issuing institution, legal character, territorial scope, relevant dates, source location, and acquisition event. It should distinguish the date of publication from the date of legal effect and from the date the item entered the corpus. Where those dates are unknown, the uncertainty should be visible. The mere fact that a file was downloaded recently does not make its contents current. Equally, older material should not automatically be discarded: it may be necessary to determine the law applicable to an earlier transaction or to understand the development of a doctrine.

Currency therefore requires relationships between records. An amendment should be linked to the affected provisions, a replacement rule to the earlier version, and a correction to the item corrected. A system designed only to select the newest file may fail when the question concerns an earlier legal period. Conversely, a system trained on historical material may confidently express a rule that no longer applies. The proposed control is version-aware retrieval and explicit dating of the answer, supported by a maintained source collection. This is an institutional responsibility, not simply a model parameter.

Judgments require comparable care. The corpus should distinguish the court's reasoning from counsel's submissions, the operative order from background discussion, and the majority position from a separate opinion. A passage can be accurately reproduced but misdescribed as a holding. An automated summary may also omit the procedural setting that limits its significance. These risks reinforce the verification protocol developed in Chapter 2: a legally consequential proposition requires inspection of the supporting authority, not confirmation that the generated sentence appears in some text.

The strongest objection is that full legal annotation is expensive and no database can be perfect. That is correct, but the alternative is not an unsupported claim of completeness. A collection can be useful within a defined scope. It should state the jurisdictions, institutions, dates, and source classes covered, identify important exclusions, and provide a referral or abstention route for questions beyond that scope. A smaller maintained corpus may be more suitable for a bounded task than a larger collection whose legal status cannot be established. Suitability must be assessed against the actual use rather than database size alone.

Language coverage and conflicting translations

Language coverage has at least three dimensions: the languages in which source material exists, the languages in which users can ask questions, and the languages in which legally significant responses are delivered. Supporting a language at the interface does not establish adequate coverage in the other two dimensions. A system may accept a local-language account, translate it into an internal working language, and retrieve authorities from a limited foreign-language collection. Each transformation can affect the representation of the legal problem.

A translation record should identify who or what produced the translation, the source version, its date, and whether the translation is official, certified, professionally reviewed, or provisional. Where several versions conflict, the system should not silently choose the most fluent wording. It should preserve the disagreement and direct the user to the applicable rule concerning authoritative language or interpretation. The chapter does not assume that one language is legally controlling across the focal jurisdictions. That status must be established for the instrument and legal setting involved.

Testing should focus on legal meaning rather than general fluency. A change from permission to obligation, an omitted exception, or an altered limitation period can change what a person does. Terms describing residence, household membership, ownership, consent, or review may not align neatly across legal and everyday language. The test collection should therefore include ordinary user descriptions as well as formal terminology. Reviewers should examine whether the translated problem retrieves the relevant material and whether the resulting response preserves the qualifications necessary for action.

Coverage claims also require appropriate denominators. A count of translated documents does not establish the proportion of relevant law translated, and a count of supported languages does not establish performance for dialects, scripts, or mixed-language inputs. The assessment should record what was tested, by whom, on which task, and with what unresolved disagreements. It should not present an untested language as equivalent to a validated one. Where adequate testing is unavailable, a restricted informational role or referral to a qualified person may be more defensible than an authoritative response.

Privacy, anonymization, and protected narratives

Public availability does not make every subsequent use harmless or lawful. A judgment can contain information about people who did not choose to become training examples. A contract may reveal business arrangements, personal circumstances, or information supplied under duties of confidence. A community account may be suitable for a limited local purpose but inappropriate for unrestricted indexing. The legal assessment must therefore examine the permission for the proposed processing in addition to the route by which the material was obtained.

Anonymization should be evaluated against the contents and context of the collection. A document may omit names while retaining a distinctive chronology, address, occupation, or family relationship. Linking several individually limited records may increase identifiability. The proposed corpus protocol consequently records both redaction methods and residual risks, including access to linking keys. It also distinguishes an authoritative original, a public redacted version, and a development copy. Mixing these versions without clear provenance can expose protected information or make the system unable to explain which text it actually used.

The relationship with judicial material requires particular precision in South Africa. POPIA's exclusion concerning judicial functions is activity-specific in its wording; it should not be treated without analysis as an exemption for any private collection containing court documents (POPIA, 2013, s. 6(1)(e)). The relevant questions concern the actor, operation, and applicable law. The same caution applies elsewhere: rules governing access to a court file, publication of a decision, reuse by a commercial service, and disclosure of a generated summary may answer different questions.

Reproduction, licensing, and training permissions

Permission should be mapped to acts rather than expressed as a single statement that a dataset is "open." Reading, downloading, reproducing, translating, indexing, training, fine-tuning, supplying material to a model provider, and redistributing content are different uses. A licence may permit some and not others. An exception may depend on the nature and purpose of the use. The rights in the underlying material may also differ from rights in annotations, translations, or the selection and arrangement of a collection. None of these questions is resolved by successful technical access.

Brazil's copyright statute illustrates the distinction. Article 8(IV) excludes specified official texts, including legislation and judicial decisions, from copyright protection under that law. Article 7 separately addresses protected works, including qualifying translations and compilations; its treatment of compilations distinguishes the arrangement from the underlying data or materials. Article 11 identifies the author as the natural person who creates the work (Copyright Law, 1998, arts. 7-8, 11). These provisions support separating the official text from added intellectual contributions. They do not, by themselves, decide every privacy, contractual, database, or training issue arising from a mixed corpus.

A legal collection assembled from official judgments and commercial commentary should therefore preserve item-level provenance and permissions. Removing a publisher's logo does not establish permission to reproduce its analysis. Paying for access does not necessarily permit supplying the collection to an external training service. Conversely, a restrictive description attached to a collection should not be assumed to establish rights over every underlying official text. The assessment should identify the particular protected contribution and the permission or exception relied

upon, rather than accept either a supplier's ownership claim or a developer's assertion of unrestricted public use.

Training raises questions at several points. The acquisition copy, preprocessing, retention of examples, model development, evaluation, and subsequent generation may each involve different material and actors. A permission adequate for a local search index may not describe a general model offered to third parties. A provider's promise concerning ownership of outputs does not establish that it had the necessary rights in inputs. The materials examined for this chapter do not resolve the legality of every commercial training arrangement; the assessment must preserve that uncertainty and obtain the relevant local legal analysis before making a definitive representation.

Generated legal content presents a related but separate issue. The institution should distinguish the source of a draft, the human contribution to its revision, contractual allocation of use rights, and the law governing any copyright claim. It should not assume either that every output is protectable or that every output is free of third-party rights. More immediately, ownership does not establish legal reliability. A firm may have permission to use a generated clause that is unsuitable for its client, and an institution may lawfully reproduce a public judgment while misrepresenting its holding. Rights clearance and professional verification remain separate tasks.

Community knowledge and the limits of inclusion

Community knowledge should not be treated as an unowned supplement to an incomplete official corpus. The assessment must ask who may provide it, what authority they possess, whose account they represent, and which restrictions attach to collection or disclosure. A recorded description of a customary practice is not necessarily an authoritative statement of that practice. It may reflect a dispute, a particular location, or the position of one participant. Translating it into a uniform rule can remove the very context that gives it significance.

The response is not to exclude non-state knowledge automatically. It is to distinguish recording, attribution, legal recognition, and permission for reuse. A corpus can preserve several accounts and identify their limitations without treating them as interchangeable with legislation. Where an application cannot accommodate contested or context-dependent material, it should state that limitation and provide an appropriate route beyond the system. This approach follows the institutional mapping in Chapter 1 while avoiding the assumption that more collection is always the remedy for exclusion.

Table 3.2 provides a corpus assessment for use by legal editors, records specialists, language reviewers, and the institution responsible for the application. It records the evidence needed to support a claim of suitability, rather than asserting that any existing collection has passed the assessment. Each finding should be linked to a versioned inventory and a responsible reviewer. Licensing and privacy restrictions should remain visible when the collection is changed or transferred.

Table 3.2. Corpus assessment for legal authority, language, privacy, and permitted use

Assessment domain	Evidence to retain	Review question	Correction or restriction
Collection scope	Inventory of jurisdictions, institutions, document types, dates, languages, and known omissions.	What legal task can this collection support, and what lies outside it?	Narrow approved use; disclose gaps; add sources only through review.
Authority and status	Official source, issuer, jurisdiction, procedural status, and distinction between holding, submission, and commentary.	Is the passage authoritative for this claim and setting?	Correct status labels; prevent unsupported authoritative presentation.
Currency and versions	Enactment, commencement, amendment, repeal, judgment, and retrieval dates as relevant.	Which text governed at the time of the matter?	Preserve historical versions; update operative text and affected outputs.
Provenance and integrity	Source identifier, original file, extraction method, alterations, redactions, and responsible editor.	Can the system answer be traced to the actual source version?	Repair omissions and extraction errors; maintain an intelligible change record.
Language coverage	Source and query languages, translation origin, authoritative status, local testing, and reviewer disagreements.	Does legal meaning survive input, retrieval, and output?	Restrict untested functions; refer contested wording for qualified review.
Privacy and confidentiality	Personal or protected content, publication restrictions, redaction method, linking keys, and residual identifiability.	Does public access authorise this particular processing or disclosure?	Limit collection and access; revise redaction; withdraw unauthorised copies.
Intellectual-property permissions	Rights-holder or exception, licence version, protected contribution, and permission for each proposed act.	Are access, copying, training, translation, and redistribution separately addressed?	Exclude unlicensed uses; renegotiate or substitute lawful material.

Table 3.2. Corpus assessment for legal authority, language, privacy, and permitted use

Assessment domain	Evidence to retain	Review question	Correction or restriction
Community knowledge	Contributor authority, consent or other permission where relevant, attribution, context, and contested accounts.	Whose account is represented, and what recognition or use limits apply?	Preserve disagreement; restrict publication or use; provide referral beyond the corpus.
Quality and suitability	Task-specific tests, temporal and linguistic gaps, source disagreement, and known failure patterns.	Does the evidence support the stated function rather than a general quality claim?	Narrow deployment; correct sources; repeat relevant validation.
Correction and release	Complaint route, editor, approval, release identifier, recipient inventory, and prior affected answers.	Can an error be repaired beyond the current search result?	Notify appropriate users; reassess earlier reliance; retain a withdrawal history.

Establish rights and legal status for each jurisdiction, source and intended act. A source can satisfy one domain while failing another.

A corpus should be approved for a defined function, population, language setting, and period, with a process for correction and withdrawal. Discovery of an omitted amendment, disputed translation, unauthorised source, or identifiable protected narrative should trigger an assessment of affected outputs. Updating the collection is only the first step when earlier answers influenced legal work. The institution must also identify whether those answers require correction. The central test is whether the system's apparent knowledge can be traced to sources whose authority, limitations, and permitted uses are understood.

3.3 Public Benefits and Identity: Historical Administrative Analogues for AI

Two functions rather than a general account of automated government

The legal significance of automated administration becomes clearer when the comparison follows a defined function. This section therefore examines two functions within Pakistan's Benazir Income Support Programme: household-based targeting of cash transfers and biometric verification when selected beneficiaries collect payments. The first concerns the identification of households for assistance. The second concerns the delivery of assistance to the intended recipient. Their location within one programme makes it possible to distinguish decisions that are often collapsed into a single description of digital welfare.

These are historical documentary cases, selected to isolate two functions rather than to measure the effects of contemporary AI. The targeting account uses the expressly identified October 2020 working-paper version of Haseeb and Vyborny's study. The payment account uses Clark and colleagues' 2022 evidence note concerning the 2017-2019 rollout. The BISP Act 2010 identifies the statutory allocation of authority. Reading these sources together permits an institutional reconstruction, while distinguishing empirical observations, statutory powers and unresolved case-specific legality.

A shared programme is a useful comparator because targeting and payment concern the same broad beneficiary relationship but operate at different stages. It holds more institutional context constant than a comparison between unrelated national projects. The design nevertheless cannot isolate the effect of AI: the documented reform includes fixed calculations, data collection, authentication technology and changes in delivery arrangements. Its contribution is to identify where a later predictive system would have to fit within an existing allocation of legal power.

Household information and benefit targeting

Haseeb and Vyborny describe a shift from politically mediated selection to household asset data and proxy means testing. Their account identifies NADRA's scoring role and the circulation of scores and eligibility status to field administration (2020, §2). Their evaluation combines difference-in-differences and close-election regression-discontinuity approaches. It supports an argument that information and restrictions on discretion changed targeting, not that every excluded household was legally ineligible or that a formula independently possessed authority to decide entitlement.

The institutional reconstruction separates four choices. Someone selects what condition the programme seeks to address. Someone chooses observable information as evidence of that condition. Someone specifies a formula and threshold. Someone uses the result to grant, refuse, continue, or stop assistance. The fact that a computer performs the calculation does not merge these choices. A poverty estimate answers a stat-

istical question; a threshold determines how that estimate is used; the final act belongs within the programme's legal and administrative framework.

An existing legal criterion can be implemented through a new evidential method without necessarily becoming a new legal rule. Conversely, an evidential method can effectively add a condition if it makes access depend on something the governing framework does not require. The distinction must be examined rather than presumed. If the legal criterion concerns resources, a database match may be one form of evidence. If the system treats absence from that database as conclusive ineligibility, it has attached a further consequence to missing information. The legal question concerns authority for that consequence, not merely the accuracy of the match.

The targeting example also shows why the role of the data custodian must be separated from the role of the deciding institution. A body that calculates a score may possess the technical ability to amend inputs or rerun the calculation. That ability does not necessarily give it authority to decide whether a household receives assistance. The programme may control the threshold but not the underlying identity record. Effective review needs both bodies' contributions without requiring the applicant to determine the allocation of technical and legal responsibility unaided.

The statutory allocation is identifiable. The BISP Act assigns approval of policies, manuals and eligibility criteria to the Board; it separately addresses delegation, management's complaints function, eligibility and the manner of disbursement (BISP Act, 2010, ss. 6(1)(d)-(e), 6(2), 10(h), 11-12). Section 21 also permits an aggrieved person or family to challenge an adverse eligibility decision within thirty days, through a representation in the form and to the authority prescribed by regulations. Accordingly, neither the scoring body's technical capacity nor a payment provider's interface supplies the relevant rule-making power. An individual historical refusal still requires the applicable Board-approved criterion and implementing instrument. The completed conclusion is institutional: the statute identifies where to locate authority, while the empirical studies identify where computation entered the process. These sources do not authorise replacing a missing implementing instrument with an assumption of either legality or illegality.

Authentication and the delivery of an existing benefit

Clark and colleagues study a package combining fingerprint authentication, personal collection by women and greater reliance on retail agents. Their difference-in-differences analysis uses 3,074 beneficiary households in 63 districts, observed in 2016 and 2019. Personal collection increased by 46 percentage points; multiple-attempt difficulties increased in the early period, while the study did not detect an overall reduction in collection over the measured period. These are effects of the rollout package, not biometrics alone (Clark et al., 2022, pp. 8-10, 14-15). The relevant legal distinction is between an entitlement decision and the practical conditions under which an already selected person can obtain payment.

What the historical comparison can establish

The payment evidence concerns beneficiaries already in the study population. It cannot by itself measure exclusion before registration or establish every person's experience of unsuccessful authentication. The note itself cautions that a question about the last cash withdrawal may fail to capture occasions on which no withdrawal occurred (Clark et al., 2022, p. 14, note 20). The appropriate conclusion is therefore about observed rollout outcomes, not a universal absence of exclusion. This limitation is substantively important because the people missing from a completed-transaction record may be those who most need an exception route.

The two functions support a bounded reconstruction of the programme's decision chain. For targeting, the institutional criterion is the Board-approved eligibility rule; household data are evidence; a computed score is an intermediate representation; and admission to the beneficiary population is the consequential act. For payment, the relevant starting point is an existing selection; a fingerprint comparison supplies authentication evidence; and successful delivery of the authorised amount is the operational outcome. The same technical word 'failure' consequently denotes different problems in the two chains.

This reconstruction produces a legal inference without pretending to determine an individual appeal. A failed biometric comparison does not amend the Board's eligibility criterion. If the institution treats it as conclusive loss of entitlement, the additional consequence requires authority and an evidential justification beyond the failure code. Conversely, the existence of entitlement does not make every proposed method of identity assurance unlawful. The distinction leaves room for necessary verification while requiring that the method and the substantive condition remain legally intelligible.

A defensible historical review would therefore request different records for the two disputes. A contested poverty score requires the contemporaneous criterion, relevant inputs, scoring procedure and available correction. A contested payment requires the selected-beneficiary record, authentication result, payment instruction, agent or transaction record and applicable exception procedure. The statute locates approval and implementation responsibilities; it does not supply a missing historical transaction. This is a completed distinction about the objects of proof, even where the evidence cannot support a finding that a named official acted unlawfully.

The comparison also cautions against transplanting an apparently successful reform. A change that improves direct control over payments may increase the burden of attendance. A targeting rule that constrains local favouritism may reduce opportunities to explain circumstances missing from the data. These are competing institutional objectives rather than a single scale of technical accuracy. The policy implication is to preserve separate measures of receipt, burden, participation and correction, and to justify the trade-offs openly. A future AI system should be evaluated against these

functions, not against a presumption that computational consistency resolves the underlying distributional choice.

Authentication asks whether presented credentials correspond to a recorded person. Eligibility asks whether the person satisfies the conditions for assistance. Payment asks whether the authorised amount reaches the intended recipient. A failure at one stage should not be silently reclassified as a conclusion at another. A failed comparison may indicate a problem with the sample, device, connection, record, or presented identity. It does not, without an additional justified inference, establish that a previously selected beneficiary has ceased to meet the programme's substantive conditions.

The same distinction applies to positive results. A successful comparison does not prove that an amount was paid correctly, that an agent made no unauthorised deduction, or that a transaction reached its intended recipient after the system recorded completion. The decision record should distinguish authentication success, payment authorisation, transaction completion, and confirmation of receipt. Combining these events into one status such as "successful" makes later investigation less informative. The chapter's proposed distinction is functional; it does not claim that the historical system retained every record described here.

A person who cannot satisfy a system's data requirement should first receive a diagnosis of the requirement. Is the missing item necessary to establish legal entitlement, necessary for one chosen verification method, or required only by the interface? An institution may lawfully need reliable identity evidence while still possessing more than one means of obtaining it. Whether an alternative is permitted must be established under the relevant programme and identity rules. The important point is that the technical failure code should not supply the legal answer by default.

What the comparison establishes and what it does not

The two functions show different relationships between information and consequence. A targeting score helps determine who enters the beneficiary population. Authentication ordinarily operates after that selection and controls access to delivery. Errors therefore have different meanings. A false household link may alter a score; a false non-match may obstruct payment to an already selected person. The appropriate correction depends on where the error arose and which institution can change the relevant record or action. A generic complaint labelled "AI problem" does not identify either.

The comparison also prevents a misleading opposition between human discretion and technological consistency. A rule-based calculation can reduce one form of discretion while leaving choices about data collection, exceptions, and assistance with documents. A human reviewer can correct an error, but can also apply an undocumented extra condition. The analytical question is which discretion the governing framework

permits, where it is exercised, and how it is recorded. Neither removing all discretion nor inserting a human signature establishes lawful administration.

The documented functions should not be relabelled as contemporary AI deployments. A fixed formula and a biometric verification service can be legally consequential without establishing the use of a particular machine-learning architecture. Where a proposed AI model predicts eligibility, identifies anomalies, or ranks files for investigation, its function must be described separately. The institution should identify the training target, permitted inputs, output, uncertainty, threshold, and actual role in the official's work. An innovation label is less informative than this account of what the system changes.

Adding AI may introduce a new point of inference even when the legal criterion remains unchanged. A model trained to predict past approvals may reproduce the decisions recorded in its training set rather than evaluate the eligibility standard itself. A model trained on investigated cases may predict the likelihood of investigation rather than fraud. These are hypothetical mechanisms requiring examination, not findings about BISP. They show why a target label needs a legal explanation. Historical institutional action is evidence of what an institution did, not automatically evidence of what it should lawfully have done.

Missing records, incorrect matches, and discriminatory effects

Missing information has several possible meanings. The relevant event may not have occurred; it may have occurred without being recorded; the record may be held elsewhere; or the system may have failed to locate it. A person may also be unable to produce a document that exists. Treating all these situations as the same negative value removes the distinction between evidence of absence and absence of evidence. The decision process should record which situation is known and what further inquiry is authorised and practicable.

Matching requires the same precision. A shared identifier can support a link, but the assessment should establish whether the link is exact, probabilistic, manually confirmed, or unresolved. Changes in spelling, household composition, address, or the date of a record may affect the interpretation. A matching threshold is a technical choice with potential administrative consequences. Its approval should identify the errors that matter and the route for resolving uncertain cases. The record should not hide uncertainty simply because the receiving application accepts only yes or no.

A statistical difference between groups should trigger investigation, not an automatic legal verdict. The analysis should examine the relevant population, missingness, case mix, decision rule, evidence quality, and consequences of error. It should then connect any alleged disadvantage to the applicable equality or administrative-law test. A disparity in attempted authentication, completed applications, or reversals may concern different stages. Combining them into a single fairness measure may obscure who was

excluded and why. The source materials used here do not establish legally adjudicated discrimination in either documented function.

An institution also needs evidence about people who never reach an operative decision. A database of completed cases can reveal some errors but not all unsuccessful attempts to register, authenticate, or obtain assistance. Evaluation should therefore examine entry failures and abandoned attempts where lawful and feasible, alongside decisions and appeals. This is a proposed measurement requirement, not a claim that every programme can immediately produce those data. Where the evidence is missing, the resulting claims about accessibility must be correspondingly limited.

Official action, urgent assistance, and continuity

The authorised official should receive enough information to distinguish an evidential problem from a substantive reason for refusal. A score without its relevant inputs, a flag without its source, or a mismatch without a failure description may be insufficient for that task. The interface should permit the official to identify uncertainty, request clarification, and use the lawful exception route. It should not require agreement with the output merely because disagreement is harder to record. Meaningful control is a working condition that must be designed and tested.

However, an exception route is not a general permission to ignore eligibility criteria. It should identify the legal power, supporting evidence, approving officer, duration, and review event. Where urgent assistance is available under the governing framework, the process should establish how it can be considered while a record is corrected. Where the institution lacks that power, it should identify the competent service or authority rather than promise relief it cannot deliver. Urgency strengthens the need for a clear route; it does not create an otherwise absent mandate.

Continuity requires separate treatment of system failure and adverse decisions. An outage may prevent authentication or retrieval without providing any evidence about the applicant. The institution should record the interruption, preserve pending requests, and activate its authorised fallback. A temporary manual process needs a way to reconcile actions when the system returns so that payments are not lost or duplicated and deadlines are not silently restarted. These are proposed operational safeguards whose legal form depends on the service, but they should be considered before the system becomes indispensable.

The strongest argument for highly automated processing is that delay, inconsistent records, and discretionary barriers in the existing process can themselves cause serious harm. That argument is important and should be tested against an actual baseline. The alternative to an unreviewable model is not necessarily an unchanged manual system. Better records, clearer eligibility instructions, staffed exception handling, and simpler calculations may address the same problem. The institution should compare feasible

arrangements by the person's ability to obtain a lawful decision and effective correction, not by transaction speed alone.

A decision record linking criteria to consequences

Table 3.3 sets out the record needed to reconstruct a consequential administrative action. It separates the governing criterion from the evidence used, the computational output from the official act, and correction of data from reconsideration of the outcome. The record should be completed for a defined case and system version. It is not an invented case history or a claim that the Pakistani programmes already maintain every field.

Table 3.3. Decision record connecting eligibility, evidence, output, and official action

Record field	Entry or evidence required	Accountable role	Question for review
Case and period	Case identifier, application date, programme, institution, location, and applicable decision period.	Programme record owner.	Which matter and legal version are being reviewed?
Legal criterion	Exact eligibility or verification rule, source, exceptions, and competent deciding office.	Legal reviewer and programme authority.	Is this an authorised criterion rather than a software requirement?
Source evidence	Fields, documents, origin, dates, reliability limits, and applicant explanation.	Source custodian and receiving officer.	What does each item prove, and what remains unknown?
Missing or disputed data	Failure type, unavailable document, contested match, alternative evidence, and pending correction.	Intake officer with defined escalation.	Was an evidential gap wrongly treated as substantive ineligibility?
System function	Fixed rule, statistical estimate, biometric comparison, or identified AI function; configuration and version.	Technical operator and system owner.	What was actually computed, and what conclusion was not established?

Table 3.3. Decision record connecting eligibility, evidence, output, and official action

Record field	Entry or evidence required	Accountable role	Question for review
Output and consequence	Score, flag, match response, uncertainty where meaningful, and proposed next step.	Operator retaining the output record.	How did this result enter the institutional action?
Officer action	Material examined, independent conclusion, departure or escalation, reasons, and authorisation.	Competent decision-maker.	Was the output assessed or simply implemented?
Notice and assistance	Decision communicated, relevant reasons, delivery channel, assistance, and applicable challenge information.	Service office and notice owner.	Could the person understand and respond to the material issue?
Urgency and continuity	Requested interim help, legal power, authorised alternative, outage record, and reconciliation plan.	Designated exception or emergency officer.	What happened while ordinary verification or processing was unavailable?
Review and restoration	Review request, preserved version, source amendment, recipient notices, reconsidered decision, and relief actually given.	Reviewing authority with source and recipient custodians.	Did correction reach the person's legal and practical position?

Original completion template. The Pakistani studies illustrate distinct historical functions; this table does not claim that all listed records or procedures existed in those programmes.

The responsible programme authority should own the final record, while source custodians attest the information they supply and technical operators identify the relevant processing. A reviewer should be able to determine whether the system applied an authorised criterion, used an unauthorised substitute, or encountered an evidential gap that was incorrectly converted into refusal. Where a person cannot satisfy the chosen data route, the record should show what alternative or review was considered. The leg-

al test remains the justified administrative act, not the successful completion of a technical transaction.

3.4 Automated Decisions in Employment, Credit, and Digital Platforms

Private decisions and the interests they affect

An automated decision need not be issued by the state to alter a person's practical legal position. A refusal to shortlist an applicant, an adjustment to a credit limit, a restriction on receiving payment, and the suspension of a platform account concern different relationships. Their common computational form does not make their governing duties identical. The appropriate legal analysis begins with the actor, relationship, affected interest, decision, and available forum. Only then should it ask how automation changes the evidence needed to establish compliance or obtain correction.

This section uses statutory illustrations rather than presenting undocumented commercial deployments as case studies. South African employment and credit legislation supplies two sector-specific settings; Brazilian data-protection, consumer, and contract legislation supplies a bounded framework for examining platform decisions. Hypothetical screening and account-restriction scenarios explain how these rules might become relevant. No particular employer, lender, or platform is said to have committed discrimination, breached a contract, or operated an unlawful model. Such findings would require evidence of an actual system and the applicable legal interpretation.

The comparison distinguishes three forms of control. A business may select the substantive criterion, a supplier may calculate the score, and an operational team may implement the consequence. These roles can coincide but should not be assumed to do so. A lender's claim that a bureau supplied the information does not identify who chose to treat it as decisive. Similarly, a platform's reference to an automated security check does not disclose whether the restriction resulted from a model, a fixed rule, or an employee's assessment. Each link needs a named actor and an intelligible record.

Automated-decision protection is not a general prohibition on AI

POPIA section 71 applies, subject to exceptions, to decisions producing legal consequences or substantially affecting a data subject when based solely on automated personal-information processing intended to provide a profile. Work performance and creditworthiness are among its examples. The contractual exception requires either that the data subject's request has been met or appropriate protection of legitimate interests; a further exception concerns decisions governed by a law or code specifying appropriate measures. For the protective-measures branch of the contractual exception, subsection (3) requires an opportunity to make representations and sufficient information about the underlying logic to enable them (POPIA, 2013, s. 71).

The applicability elements in section 71(1) must be satisfied together. Section 71(2), however, provides alternative exceptions, and subsection (3)'s safeguards attach to the specified protective-measures branch. Neither 'automation is allowed with consent' nor 'every AI decision requires human approval' accurately describes this structure. A decision outside section 71 can still engage other processing conditions and sector duties. Substantive human review therefore changes one part of the legal analysis without erasing the rest.

Meaningful review is therefore both a legal and an operational inquiry. A system may require a human click without permitting the person to inspect the evidence, correct a field, change the result, or record a different reason. The research should establish those conditions before describing the decision as substantively reviewed. Conversely, the mere use of a score does not establish that the decision was solely automated. The chapter proposes examining actual discretion, available information, review time, and recorded action rather than deciding the issue from an interface label.

Hiring, assessment, and workplace monitoring

For employment within its scope, South Africa's Employment Equity Act prohibits direct and indirect unfair discrimination in employment policies and practices. Applicants are included for the protections in sections 6-8. Psychological testing and similar assessments engage the requirements of scientific validity, reliability, fair application, and freedom from bias in section 8(a)-(c). Whether a particular digital assessment is such a test depends on its function, not its marketing name. The 2013 and 2022 amendments must be read with the principal Act; the later amendment removes the certification paragraph previously added to section 8, not its validity and fairness requirements (Employment Equity Act, 1998, ss. 6, 8-9; Employment Equity Amendment Act, 2013, ss. 3-4; Employment Equity Amendment Act, 2022, s. 2). The 2022 amending Act took effect on 1 January 2025 (South African Government, 2024, commencement announcement). Historical analysis must also account for the High Court's earlier order invalidating commencement of the section 8 amendment and preserving the preceding provision (Association of Test Publishers v President, 2017, paras. 37-38).

A hypothetical hiring model illustrates the evidential questions. Suppose an employer predicts retention using employment gaps, travel distance, and earlier managers' performance ratings. The first inquiry is what retention means and whether it is an appropriate selection objective for the post. The second concerns the evidence: which workers were included, how departures were recorded, and whether the ratings measure the stated capability. The third follows the output into the decision. A weak recommendation that invites further discussion differs from a threshold that removes an applicant before anyone examines qualifications.

Removing explicit protected characteristics does not end the inquiry. A variable may be associated with a protected characteristic, an employment opportunity, or a legit-

imate task requirement, and those possibilities require different evidence. The assessor should investigate the variable's meaning, its relationship to the job, the consequences of its use, and available alternatives. A statistical association is a reason to investigate; it is not, without the relevant legal test and facts, an established finding of unfair discrimination. The applicable burden of proof also matters: section 11 distinguishes allegations involving listed grounds from allegations on arbitrary grounds (Employment Equity Act, 1998, s. 11, as substituted by Employment Equity Amendment Act, 2013, s. 6).

Monitoring introduces a temporal problem that an entry-stage assessment can miss. In a constructed example, data collected to verify arrival at a workplace are later used to infer commitment or future misconduct. Accuracy of the arrival record would not resolve the lawfulness or validity of that inference. The data-use register must identify the changed purpose, the additional attributes produced, and the consequences attached to them. A notice stating that monitoring occurs cannot, by itself, explain why every subsequent use is necessary or justified. POPIA's purpose, further-processing, and quality conditions remain relevant where the Act applies (POPIA, 2013, ss. 10-16, 18).

Review should distinguish an inaccurate event record from a contested assessment of its significance. An employee might accept that a device recorded inactivity while explaining that the task required work away from that device. Correcting the record is not the only possible response; the interpretation, metric, or decision may need reconsideration. The proposed employment review file therefore preserves the task description, monitored event, inference, relevant explanation, decision rule, and accountable manager's response. It does not assume that greater measurement necessarily provides better evidence of work.

Credit assessment, reasons, and correction

Credit decision-making provides an example in which general data rights intersect with specific sector obligations. For agreements within its scope, South Africa's National Credit Act addresses unfair discrimination, access to reasons, credit-information challenges, and responsible assessment. Section 62 requires a written statement of the dominant reason for specified adverse credit decisions when the consumer requests it. Section 72 provides routes to inspect and challenge credit information. Section 82 allows providers to choose assessment models, but requires fair and objective assessment consistent with the applicable affordability regulations (National Credit Act, 2005, ss. 61-62, 72, 81-82; National Credit Amendment Act, 2014, s. 24).

The distinction between a dominant reason and a technical description is important. “The system assigned a low score” identifies an output, not necessarily the decisive reason for refusal. A useful account should connect the adverse decision to the relevant information and assessment criterion without pretending that every possible internal parameter explains the individual result. The information needed to contest a misre-

ported debt may differ from that needed to challenge an income estimate or an inappropriate model application. The legal duty and the practical explanation should be examined together, not replaced by a generic demand for source code.

A hypothetical lender using an applicant's transaction record illustrates the difference between prediction and affordability. A model may estimate the likelihood of repayment from patterns among previous borrowers. The legally relevant assessment may also require attention to existing obligations and the financial position of the particular consumer. A predictive measure cannot be assumed to answer every statutory question. The analytical record should state which obligation the model supports, what remains for other evidence, and whether missing records lead to additional inquiry or automatic refusal. No universal entitlement to a loan follows from the existence of a review route.

Correction may require action by more than one body. The bureau can examine a disputed entry; the provider can reconsider the assessment that relied upon it. A data-protection complaint, a challenge to credit information, and a complaint about the credit provider's conduct are not interchangeable. The available routes include the Information Regulator under POPIA and the National Credit Regulator under the credit statute, according to the allegation and their respective powers. The institution should explain the distinction while preserving any other competent forum rather than send the person repeatedly between agencies (POPIA, 2013, ss. 24, 74-75; National Credit Act, 2005, ss. 72, 136).

Platform restrictions, standard terms, and contested status

Platform participation may involve several simultaneous relationships: a worker providing services, a merchant accessing customers, a consumer using an application, or a person receiving payment through an intermediary. Employment classification and data protection answer different questions. A dispute about employee status does not itself displace a data right whose statutory conditions are independently met. POPIA's scope, for example, turns on its definitions and application provisions, not solely on whether the data subject is an employee (POPIA, 2013, ss. 1, 3). The decision record should therefore identify each relationship and legal route separately.

Brazil's LGPD offers a useful comparison. Article 20 provides a right to request review of solely automated personal-data decisions affecting the person's interests, including professional and credit profiles. The current wording does not retain the original express reference to review by a natural person. Paragraphs 1-2 address information about decision criteria and procedures, commercial and industrial secrecy, and a possible ANPD audit of discriminatory aspects where information is withheld on secrecy grounds. These provisions should not be converted into either an unrestricted source-code right or a finding that any automated review is necessarily adequate (LGPD, 2018, art. 20, as amended).

Consider a hypothetical platform that restricts an account after detecting an apparent identity mismatch. An effective inquiry would distinguish account authentication, suspicion of misuse, contractual grounds for suspension, retention of earned payments, and permanent termination. The same initial flag need not justify every consequence. The person may need to correct an identity record, contest a match, challenge the interpretation of a term, or establish an entitlement to money already earned. An appeal that merely reruns the same input through the same configuration would not investigate all these questions. The chapter proposes an appropriately authorised reviewer and a record of the issue actually reconsidered.

Contract law supplies a further, distinct inquiry. Brazil's Civil Code requires good faith in concluding and performing contracts, while also expressing limits on contractual revision and recognising agreed risk allocation within its framework. For adhesion contracts, articles 423-424 address ambiguous terms and certain advance waivers. These provisions require analysis of the actual agreement and relationship; they do not establish that every account restriction is invalid or that every standard term is enforceable (Civil Code, 2002, arts. 421-424). The relevant version of the terms, notice of change, stated grounds, and action taken should therefore be preserved alongside the system output.

Consumer protection must likewise be applied through its threshold conditions. Brazil's Consumer Defence Code defines the consumer and supplier relationship and provides for clear information and control of abusive terms. Its definition of services expressly distinguishes employment relationships. The chapter consequently does not assume that every worker or merchant using a platform is a consumer of that platform for every disputed act. Where the Code applies, information and standard-term obligations may be relevant in addition to LGPD rights; where it does not, the analysis must identify the applicable contractual, labour, or other route (Consumer Defence Code, 1990, arts. 2-3, 6, 51).

Evidence, review, and the limits of common protocols

Commercial confidentiality and fraud prevention present a serious counterargument to extensive disclosure. Revealing operational controls can create risks, and institutions may need to protect other people's information. The appropriate response is a differentiated evidence arrangement: intelligible reasons and relevant personal information for the affected person, protected technical access for a competent reviewer where authorised, and documented grounds for withholding particular material. This is a proposed implementation approach, not a claim that every jurisdiction already provides the same disclosure entitlement. Confidentiality should be analysed specifically rather than used as a universal answer to any request for explanation.

Review also has an economic dimension. A procedure can be formally available while requiring repeated submissions, expensive assistance, or continued loss of income during investigation. The proposed evaluation records time to acknowledgment, time

to substantive consideration, evidence requests, outcome, and restoration actually provided. It should distinguish a complaint closed in the database from a disputed payment released or an erroneous restriction removed. Temporary protective action should be considered where available under the applicable law and contract, with reasons for granting or refusing it.

Table 3.4 organizes the sector comparison without supplying findings about an unnamed system. Its source references identify statutory starting points, not exhaustive advice on jurisdiction, limitation periods, or litigation strategy. The named decision-maker must be replaced with the actual legal entity. The evidence column is a proposed investigation record rather than a statement that every item is automatically disclosable to every complainant.

Table 3.4. Sector comparison of private automated decisions and routes of protection

Decision, actor, and interest	Relevant duty or legal question	Evidence to obtain	Complaint or review route
Hiring or work assessment, South Africa. Employer and any assessment provider; access to employment or treatment at work.	Employment Equity Act, ss. 6, 8-11; applicable POPIA processing conditions and s. 71. Establish employment/applicant coverage and the actual assessment function.	Job requirements; features and labels; validation; affected population; selection threshold; actual reviewer action; relevant explanations.	Employer process; CCMA conciliation and the applicable Labour Court or arbitration route under s. 10; Information Regulator for qualifying data complaints (POPIA, ss. 74-75).
Credit assessment, South Africa. Credit provider and relevant bureau; finance, terms, and accuracy of credit information.	National Credit Act, ss. 61-62, 72, 81-82, within scope; POPIA where applicable. Distinguish repayment prediction, affordability assessment, and reasons.	Application and financial evidence; bureau source; model and rule version; dominant reason; disputed entries; effect of correction.	Provider and bureau challenge processes; National Credit Regulator (s. 136); Information Regulator for data issues; other competent statutory forum as applicable.

Table 3.4. Sector comparison of private automated decisions and routes of protection

Decision, actor, and interest	Relevant duty or legal question	Evidence to obtain	Complaint or review route
Solely automated platform restriction, Brazil. Actual controller and operating entity; account participation, professional profile, or access to payment.	LGPD, arts. 18, 20-22. Establish the personal-data contribution, affected interest, and decisional role. Article 20 does not itself retain an express natural-person-review requirement.	Input, flag, decision criteria, terms version, restriction type, review request and response; evidence relevant to any secrecy claim.	Request to controller; petition to ANPD under art. 18; competent court under art. 22. The availability and scope of other routes require separate analysis.
Platform contract or consumer restriction, Brazil. Contracting entity and affected user; contractual performance, information, or consumer interests.	Civil Code, arts. 421-424; Consumer Defence Code, arts. 2-3, 6, 51, where the relationship qualifies. Do not assume worker, merchant, and consumer status coincide.	Applicable agreement; notice of change; stated ground; payment and service record; disputed term; status and relationship evidence.	Contractual complaint process and competent civil court; consumer-protection bodies where applicable. Employment jurisdiction depends on the actual relationship and governing law.

Read the cited provisions with applicable amendments and procedures. The evidence column identifies the inquiry to undertake; access to the records requires its own legal route.

The responsible institution should retain separate conclusions about data lawfulness, sector compliance, contractual authority, and the operation of review. An employment-status dispute should not prevent examination of an independently applicable data right. An accurate score should not settle whether a particular consequence was permitted. A successful contractual claim should not be presumed to resolve every downstream record. The chapter's central requirement is attribution: each material use of information and each decision should remain connected to an actor, a duty, evidence, and a competent route of challenge.

3.5 Cross-Border Processing and Technology Supply Arrangements

A locally delivered service may depend on several foreign actors

A person may submit an application to a local institution while the information passes through a hosting company, a model provider, a logging service, and a remote support team. The outward-facing service name does not reveal these relationships. For legal analysis, the relevant questions concern which actor receives which information, what that actor may do, where access occurs, and whether the institution can obtain evidence and maintain its function when an arrangement changes. A list of suppliers is insufficient unless it is connected to actual processing and decision dependencies.

A hypothetical cloud-based legal service illustrates the distinction. Its database is stored in the institution's country, but selected passages are sent to a foreign model endpoint to produce an answer. A separate provider records prompts for troubleshooting, and an overseas support team can inspect exceptional transactions. A statement that the database is locally hosted would describe only one part of the arrangement. Whether each event constitutes a regulated transfer, disclosure, or other processing activity requires the applicable legal analysis. The map should nevertheless disclose each event before the institution relies on a general claim of local storage.

Four locations should be distinguished: the legal establishment of an actor, the place where records are stored, the place from which access is exercised, and the forum in which an obligation can be enforced. They may differ. Technical routing information does not itself establish legal jurisdiction, just as a contractual governing-law clause does not physically relocate a server. The assessment must connect the actual architecture to the particular rules rather than infer the legal position from the supplier's headquarters or the selected cloud-region label.

The dependency register should include service data, model inputs, outputs, diagnostic records, account identifiers, access credentials, backups, and derived stores where relevant. Information excluded from the main database can still appear in a prompt or support ticket. The amount retained should be proportionate to the purpose, but the existence of these copies must not be concealed by treating them as incidental. In an incident, a supposedly peripheral record may be the evidence needed to establish what was disclosed or which version produced a disputed result.

Transfer permission is separate from the basis for processing

POPIA section 72 does not impose a general rule that all personal information must remain in South Africa. It permits transfers to a foreign third party through specified routes. These include protection through a law, binding corporate rules, or a binding agreement satisfying the section's adequacy requirements, including onward-transfer protections. Other routes concern consent, necessary contractual arrangements, or a qualifying benefit to the data subject where consent is not reasonably practicable and would likely be given. Each route has conditions; the existence of a foreign service contract alone does not establish that they are met (POPIA, 2013, s. 72).

The analytical consequence is that the institution needs two connected findings. First, the proposed processing must be lawful for the underlying task. Second, any regulated foreign transfer must satisfy an available transfer route. A consent record directed to delivery of a service should not be assumed to cover foreign model development. Equally, an adequate recipient arrangement does not cure excessive collection, incompatible reuse, or an unauthorised substantive decision. The data-use register in section 3.1 and the dependency register here should therefore refer to one another without merging their separate conclusions.

Prior authorisation may add another condition. Section 57 identifies specified processing activities, including certain new-purpose linkage using unique identifiers, credit reporting, and transfers of special personal information or children's information to a foreign third party without adequate protection. The section also contains a qualification involving an applicable code of conduct in force. An institution should examine that framework and the related procedure rather than conclude that a section 72 route disposes of every pre-deployment requirement (POPIA, 2013, ss. 57-58). These provisions are not a general licensing scheme for all AI systems.

Brazil's operative transfer instrument is Resolution CD/ANPD 19/2024, read with its annexes and the published 2025 correction. Annex I distinguishes international collection from an international transfer and requires both a processing basis and a valid transfer mechanism. The standard-clause route requires the prescribed text without alteration; bespoke specific clauses require approval under a different procedure (ANPD Resolution 19/2024, Annex I, arts. 3-9, 15-16, 21-24). An imported contractual label, or a clause accepted in another jurisdiction, does not establish which Brazilian mechanism the parties have satisfied.

The comparison concerns two ways of preserving protection when processing crosses a border. POPIA section 72 asks whether an identified statutory route and its conditions are met. Brazil's regulated standard-clause route supplies a prescribed contractual architecture. Neither route answers the underlying question whether a court or legal service may disclose the particular material. Transfer compliance therefore follows, rather than replaces, identification of the actor's authority, confidentiality obligations

and processing purpose. The following worked flow applies the distinction to a proposed service rather than certify an unnamed supplier.

Applying the transfer rules to a defined service

Consider a Brazilian legal-information provider proposing to send a user's account to a foreign model provider for summarisation. The database stays in Brazil, but the receiving company can access the transmitted text. The analysis starts with a map of that disclosure rather than the database location. If the user instead supplies information directly to a foreign service, Annex I's distinction concerning international collection becomes relevant. Neither arrangement should be chosen merely to obtain a convenient label: the factual allocation of collection, access and purpose determines which question must be answered.

Assume for this exercise that the parties choose the Brazilian standard-clause mechanism. The implementation record must identify the exporter and importer, the data and purpose, and the security arrangements, and must preserve the prescribed text. A supplier's general promise to comply with applicable law is not equivalent to executing the required instrument. Under Annex I article 17, the controller must provide the clauses on request within fifteen days unless ANPD sets another period, subject to commercial and industrial secrecy, and publish accessible Portuguese-language information about the transfer. These are legal requirements, not optional fields invented for this book (ANPD Resolution 19/2024, Annex I, arts. 16-17).

Suppose the supplier proposes to reuse the account to improve a general service for unrelated customers. This is a change in purpose and recipients' interests, not merely a technical incident within the original request. The fact that a transfer mechanism exists does not resolve the legality of that additional processing. The institution must either establish a proper basis and compatible arrangement or exclude the reuse. The worked conclusion is therefore conditional approval of a defined flow, not approval of the supplier as an entity for every future activity.

The record must also survive correction. If a material date is wrong, changing the Brazilian source file does not show that the foreign copy or a later summary has changed. The prescribed clauses address notification of correction and related measures to recipients, subject to stated exceptions, and mutual assistance with data-subject requests (Annex II, cls. 15.5-15.6). The practical completion evidence is an identified correction event, the relevant recipients and their response. Where an earlier output influenced advice or another decision, a separate professional or institutional review remains necessary.

Foreign access demands illustrate why a contract is useful without being sovereign. Annex II clause 19 provides a framework for notice, legally grounded challenge and records of access requests, while recognising that destination law can bar notice. This does not guarantee that the exporter will always obtain immediate warning. An insti-

tution handling sensitive legal material must therefore evaluate whether its function can tolerate the residual constraint. Contractual protection is a relevant safeguard; it is not evidence that a foreign public authority is bound by the parties' promises.

The South African comparator asks the same functional questions but follows its own legal route. The responsible party identifies the underlying processing justification, determines whether section 72 applies, and documents the selected route and its conditions. It separately examines any section 57 prior-authorisation requirement. The common documentation can record purpose, recipient, onward access, correction and exit; the legal conclusion must remain jurisdiction-specific. This permits interoperability of records without asserting equivalence of the two transfer regimes.

A change in supplier or support location reopens the affected part of the assessment. It should not require rewriting a record whose facts remain correct, but it must not inherit an earlier conclusion without checking the new flow. The decisive practical test is whether the institution can continue lawful processing, answer a complaint and reconstruct a consequential output after the change. A contract that permits export of data but omits the metadata needed to interpret it may provide formal possession without usable continuity.

Contracts, operators, and security responsibilities

POPIA illustrates how a statutory duty and a supplier agreement can operate together. Section 21 requires a written contract through which the responsible party ensures that an operator establishes and maintains the security measures contemplated in section 19. Section 20 addresses authorised processing and confidentiality. The contract is therefore a mechanism for implementing particular obligations, not evidence that responsibility disappears when processing is outsourced (POPIA, 2013, ss. 19-21). Whether a supplier is acting only as an operator or determines a separate processing purpose must be examined from the arrangement rather than presumed from the chosen label.

Contractual duties should follow material subcontracting. The institution should identify who may appoint another provider, what information the provider receives, whether it can use that information independently, and how the original institution can secure assistance. A promise made by the immediate supplier may be operationally weak when an essential subprovider is neither identified nor required to preserve relevant evidence. The proposed contractual response is a documented chain of equivalent obligations, notice of material changes, an objection or exit procedure where appropriate, and a duty to assist with lawful requests. These are design recommendations requiring adaptation, not a claim that every term is expressly prescribed by POPIA.

Security notification also requires precision. Under section 21(2), an operator must notify the responsible party immediately when there are reasonable grounds to believe

that personal information has been accessed or acquired by an unauthorised person. Section 22 governs notification by the responsible party to the Regulator and identifiable data subjects as soon as reasonably possible, with specified considerations and qualifications. The chapter does not import a universal seventy-two-hour period from another legal system (POPIA, 2013, ss. 21-22). A contract may need faster internal reporting to enable compliance, but contractual notice and statutory notification remain different obligations.

Not every consequential incident is a confidentiality breach. A corrupted translation, an incorrect identity match, a silent threshold change, and a prolonged outage may require different responses. Conversely, a service can remain available while disclosing information improperly. The proposed incident scheme therefore distinguishes security, integrity, model-performance, procedural, and continuity events. For each, the institution should specify who receives the alert, what evidence is preserved, whether reliance must pause, and which legal reporting or affected-person communication requirements apply. One generic incident label should not obscure the difference between restoring a server and reconsidering decisions.

Localization, control, and the capacity to obtain evidence

Local storage may address a particular transfer restriction or reduce a particular dependency, but its consequences must be established rather than assumed. It does not, by itself, prove that local officials can inspect a proprietary model, that remote support lacks access, or that a usable backup exists. Equally, foreign processing is not automatically incompatible with institutional control where applicable law permits it and adequate arrangements are demonstrated. The relevant comparison concerns the particular data, legal restrictions, security capabilities, evidence rights, cost, and continuity needs, not an abstract preference for one location.

A rigorous comparison would ask the same questions of local and foreign options. Who holds the encryption keys? Who can change access permissions? Can the institution obtain records without purchasing a new service? Are administrators and reviewers trained to interpret them? Can the system be suspended without losing access to the decision file? Where a jurisdiction imposes localization, that rule must be followed within its scope. The other controls remain relevant because complying with a storage requirement does not establish that the entire service is lawful or reviewable.

Evidence held abroad presents several distinct legal problems. Territorial execution of a warrant, an order directed to a person within jurisdiction, contractual production, voluntary disclosure and international cooperation are different routes. Their availability depends on the instrument, addressee and control of the records. A supplier's foreign address is therefore insufficient to establish either that evidence is beyond reach or that domestic process can obtain it. The dependency assessment must identify the route actually available and the time within which it can preserve effective review.

The proposed assessment identifies the body seeking evidence, the legal power invoked, the person to whom the demand would be directed, the location and control of the records, and any conflicting obligation. It should separately record whether cooperation or another lawful procedure is needed, the time required, and the risk that the evidence will be lost meanwhile. The institution should not assume that a domestic contract binds a foreign public authority or that an international cooperation mechanism will operate within an urgent service deadline. Neither should it treat uncertainty as proof that access is impossible.

Contractual evidence access is consequently valuable but limited. It can require the supplier to preserve and provide relevant material, assist authorised reviewers, explain records, and support lawful proceedings. It cannot create a statutory inspection power or guarantee that another jurisdiction will enforce every requested act. The contract should disclose known restrictions and provide an escalation process when production is contested. For critical functions, the institution should avoid making reconstruction wholly dependent on a future discretionary response from an actor over whom its practical control is uncertain.

Records needed to investigate and reconsider

An investigation requires a bounded reconstruction of the event, not indiscriminate collection of everything the system has ever processed. The proposed minimum includes the case or transaction identifier, relevant input snapshot or verifiable reference, source and time, applicable rule and model version, configuration, output, human action, resulting decision, and subsequent correction. Access and security records are added where they bear on the event. The record should preserve both what the system produced and what the institution actually did; a generated recommendation and a final decision are not the same evidence.

Reconstruction does not always require possession of the entire model or its training corpus. The required evidence depends on the allegation and the available legal powers. A wrong source field may be demonstrated through the authoritative record and input log. A disputed model change may require version history, validation material, and technical assistance. A systemic discrimination allegation may require broader protected analysis. The proposed documentation schedule should identify these levels of access in advance rather than promise unlimited disclosure that the institution cannot lawfully or practically deliver.

Evidence preservation must remain compatible with retention and security duties. Relevant records should be protected against alteration, assigned reliable timestamps, and made interpretable through schema and version information. Access should be limited and recorded. Where an investigation or proceeding requires preservation, the institution should document the basis and scope rather than silently retain every unrelated record. When a record is corrected, the earlier version may need to remain avail-

able as evidence of the disputed decision while being excluded from further operational reliance. Preservation and continued use must therefore be managed separately.

The test is whether a competent reviewer can follow the path from information to consequence. A supplier's assurance that the model was functioning normally does not establish that the input was correct, the category lawful, or the decision authorised. Similarly, proving that a foreign server was secure does not establish that a person received the correct payment. The documentation pack should permit distinct inquiries into data handling, system behaviour, institutional action, and the adequacy of correction.

Updates, outages, supplier replacement, and exit

A dependency can change without a new contract title. A provider may replace the model behind an endpoint, introduce a new diagnostic service, move support, change a retention setting, or withdraw an older version. The institution should define material changes by their effects on purpose, data access, performance, legal consequences, and evidence availability. Routine maintenance and material substitution need not follow identical approval procedures, but both should leave a usable record. Emergency security changes may require an expedited process followed by documented review; urgency should not erase the change history.

Version control must connect to decisions already taken. Where a correction or update changes an output, the institution should be able to identify affected cases and decide whether they require reconsideration. Rollback is not automatically sufficient: an older version may itself be defective or insecure. The proposed change record therefore states the reason, affected components, evaluation, approval, permissible use, and remedial action for earlier decisions. The dependency register supplies the actors and evidence needed to perform those tasks.

Continuity arrangements should preserve the legal function rather than promise only technical uptime. A manual route may be needed for urgent applications; a restricted service may be preferable to issuing unreliable results. The design should identify which activities can pause, which require alternative handling, and who is authorised to make that determination. It should also explain how records created during an outage will be reconciled without duplicate payments, lost filings, or inconsistent decisions. These are foreseeable workflow questions, not proof that a particular provider will fail.

Exit requires more than a download button. The institution needs readable data and metadata, necessary configurations and licences, unresolved complaint files, decision histories, correction records, and assistance for a successor where contractually secured. Export testing should establish that identifiers, meanings, versions, and access restrictions survive the move. The retention schedule must cover both returned material and residual supplier copies. Deletion should be verified to the extent practicable,

with backups, legal holds, and technical limitations identified rather than hidden behind an absolute promise.

Table 3.5 records these dependencies as proposed assessment fields. It is to be completed for each materially distinct actor or flow, not once for an entire product. The legal reviewer supplies the actual permission and qualifications; the operational and technical owners supply evidence of control, access, and continuity. Unknown locations, undeclared processing purposes, and unavailable essential records should remain visible as unresolved matters.

The completed register should show whether the institution can investigate an incident, correct consequential decisions, and preserve those responsibilities through supplier change. Transfer permission, local storage, security, and contractual access each answer only part of that inquiry; none replaces the others.

Table 3.5. Cross-border dependency, evidence-access, and continuity register

Actor, location, and function	Permissions and limits	Evidence and access rights	Continuity and change obligations
Institution / responsible party. Identify legal entity, establishment, function, and accountable office.	Underlying processing authority; permitted purposes; actual transfer route; relevant sector and confidentiality rules.	Local index linking activity, dataset, case, model version, output, decision, and correction records.	Named stop authority; funded fallback; control of approvals, complaints, and supplier exit.
Primary hosting provider. Record storage, backup, recovery, and administrative-access locations.	Authorized data classes and uses; relevant transfer basis; retention, security, and subcontracting restrictions.	Exportable records; access logs; service and incident evidence; defined institutional and authorised-reviewer access.	Tested restoration and export; notice of region changes; return and verified disposal subject to lawful holds.
Model provider. Identify endpoint operator, legal entity, model, version, and processing locations.	Permitted inputs; limits on provider training, retention, diagnostic reuse, and onward disclosure.	Relevant input/output record, configuration history, version identification, validation information, and investigation support.	Notice of substitution or withdrawal; approved changes; rollback or alternative-service plan.

Table 3.5. Cross-border dependency, evidence-access, and continuity register

Actor, location, and function	Permissions and limits	Evidence and access rights	Continuity and change obligations
Retrieval or derived-data service. Identify index, embedding store, cache, and source locations.	Source licences and processing permissions; restrictions on protected material; retention and correction duties.	Source-version links, ingestion and removal history, export format, and correction acknowledgments.	Rebuild or migration capability; prompt propagation of source changes; disposition of obsolete stores.
Support and diagnostic providers. Record remote-access countries and actual support roles.	Access purpose, approved personnel, confidentiality, data minimisation, and secondary-use limits.	Tickets, approvals, accessed records, incident communications, and relevant access history.	Escalation contacts; reporting duties; revocation of access; continuity during personnel or provider change.
Subprocessors and onward recipients. Identify each entity and the function performed.	Onward-transfer conditions; contractual flow-down; approval requirements; independently determined purposes.	Recipient register, relevant agreements, access evidence, preservation and lawful-production assistance.	Notice of replacement; objection or exit process where agreed; continued evidence access and deletion duties.
Institutional evidence repository. Record custodian, storage, access tiers, and preservation purpose.	Applicable retention schedule, investigation or litigation hold, and confidentiality restrictions.	Decision reconstruction, integrity checks, timestamps, schema, version history, and chain of custody as needed.	Independent access during outage or dispute; restoration tests; controlled release and disposal.
Successor or exit service. Identify transition actors, locations, and open obligations.	Rights to use transferred material, licences, permitted migration, and any new processing or transfer assessment.	Tested data/metadata export; unresolved cases; records of return, residual copies, holds, and deletion verification.	Transition assistance, reconciliation, access revocation, acceptance testing, and post-exit complaint support.

Proposed actor-by-actor register. Unknown locations or unavailable essential evidence are unresolved findings, not implied permissions. Legal production powers, contractual rights, and practical retrieval capability must be recorded separately.

Chapter 3 Key Arguments

Data governance distributes authority across a sequence of acts. Collection, inference, transfer, decision and correction require separate legal conclusions even when they share a technical system. The chapter's contribution is to connect those conclusions through records that identify the criterion applied, evidence used, actor responsible and route of challenge. This makes the distinction between a defective record and an unjustified decision operationally usable.

The historical BISP comparison separates admission to a beneficiary population from authentication and delivery. Its statutory and empirical sources establish the relevant institutional functions without resolving every historical refusal. The employment, credit and platform examples then show why a shared computational form does not create a shared legal test: processing rights, sector duties and contractual powers can overlap while retaining different thresholds and remedies.

The South African and Brazilian transfer analysis reaches a parallel conclusion. A common record can make a data flow reconstructable, but the lawful route depends on the governing instrument. Prescribed clauses, an adequacy route, confidentiality safeguards and evidence access perform different work. Their combination should be evaluated against the institution's ability to investigate and correct, rather than treated as a cumulative score in which one strength cancels another failure.

The five instruments use a common set of identifiers for the activity, case, source record, dataset, model or rule version, recipient and corrective event. Not every task needs every field. A bounded public-source search needs less information than a disputed inference transferred to an external provider. Additional documentation is justified by the decision or challenge it supports. The completed exercise in §2.5 illustrates this proportional approach; Chapter 4 applies the resulting evidence to liability and relief.

CHAPTER

04

Liability, Evidence, and Legal Remedies

IN THIS CHAPTER

4.1 Allocation of Responsibility Across AI Systems

4.2 Notice, Reasons, and Access to Decision Records

4.3 AI-Related Evidence and Standards of Proof

4.4 Individual Review and Corrective Remedies

4.5 Collective Harm and Institutional Enforcement

Introduction

An AI-related dispute does not begin with a complete account of the technology. It may begin with a payment that has stopped, an unexplained restriction on an account, a rejected application, or a document whose authenticity is contested. The affected person may know the consequence without knowing whether a model, an ordinary software rule, an inaccurate record, or an official's interpretation produced it. A practical account of remedies must therefore start with the adverse act and the legal interest involved. It should not require the person to identify a model defect before an institution will consider whether its decision was justified.

The central argument of this chapter is that responsibility, evidence, and remedy must be connected without being collapsed into one question. Identifying who controlled a technical feature does not, by itself, establish liability. Establishing that a record is authentic does not prove that its contents are true. Correcting an inaccurate field does not necessarily restore the payment or opportunity lost because of it. Conversely, an institution may be able to reconsider a decision before a court has determined which supplier, official, or other participant must bear the resulting financial loss. The different questions require different evidence, competent bodies, and legal tests.

Colombia's Constitutional Court judgment T-323/24 provides a bounded starting point. In that case, the Court examined a judge's use of ChatGPT in a tutela proceeding. It found that the judge had reached and justified the decision independently before incorporating the tool's responses; the AI use therefore did not invalidate the judgment on the ground of substitution of judicial reasoning. The Court nevertheless identified shortcomings in transparency and responsibility and explained requirements concerning verification, explanation of AI use, and privacy. It also directed the development of institutional guidance and training (Corte Constitucional de Colombia, 2024, paras. 359-376; operative orders five and six). These conclusions concern the Colombian proceeding and the Court's reasoning. They are not a universal permission to use generative AI in adjudication or a finding about a model supplier's civil liability.

The comparative significance is methodological. The same event can raise a question about the validity of a particular decision, a question about the conduct of the person using a tool, and a question about the safeguards required for future institutional use. A satisfactory response need not be identical across those dimensions. This chapter consequently separates investigation of the event from the legal characterization of conduct and from selection of relief. It asks what must be established for each proposed intervention, rather than assuming that the label “AI failure” answers all three inquiries.

The chapter uses selected functional comparisons rather than a complete survey of the six focal jurisdictions. South Africa and Brazil supply a paired examination of false matching, processing duties and remedial routes; Indian and Pakistani evidence rules test different foundations for using electronic records. Kenyan standing provisions and

Colombia's tutela procedure identify additional institutional alternatives. These comparisons hold the practical problem constant where possible while allowing the governing cause of action, procedure and institution to differ. They do not establish comparative national performance.

The method combines close reading of identified legal provisions with constructed applications. The false-match case developed below is an analytical device, not a reported incident or an estimate of prevalence. Its stipulated facts allow the inquiry to distinguish what follows from a processing error, what additionally requires proof of loss, and what turns on the public or private character of the decision. The protocols consolidate the analysis into adaptable records; their legal content comes from the relevant domestic route, not from the form itself.

4.1 Allocation of Responsibility Across AI Systems

Control, conduct, and the legal basis of responsibility

The question "Who controlled the feature that contributed to the harm?" is a useful starting point because it directs attention away from the system's public name and towards particular acts. The relevant feature might be a matching threshold, a dataset label, a restriction on human review, a rule for deleting logs, or a supplier update. Control must be established for the feature and period in dispute. A company that developed an earlier model may not have selected the local decision threshold. An institution that cannot alter the model's weights may nevertheless control whether its output blocks a payment.

The next question is legal rather than technical: which rule connects the identified conduct to responsibility? A person might possess the ability to change a system without owing the particular duty alleged by the claimant. Another actor might be legally responsible for a decision even though it cannot modify the supplier's software. The investigation should therefore record control, conduct, duty, breach, causal contribution, and the claimed consequence separately. Combining them in an assertion that the "AI owner" is responsible makes it difficult to identify either a defensible claim or the evidence necessary to answer it.

Control is relevant in different ways across legal regimes. It may identify a statutory processing role, explain who could prevent a mistake, locate documents, or establish what a contracting party promised. None of these connections alone proves a general duty to compensate. The chapter therefore treats the lifecycle map as the first stage of attribution and the domestic cause of action as the second. A claimant's inability to identify the model developer need not prevent a challenge to an institution's own consequential act.

Following responsibility through the lifecycle

At development, the inquiry concerns the function designed, the uses represented as appropriate, the limitations known, and the evidence supporting the supplier's claims. The relevant conduct may involve a representation about performance, an omitted limitation, or a change made without an agreed approval. None is proved simply because an output was wrong. The investigator needs the specifications, test conditions, release notes, communications, and actual version supplied. A model's later behaviour should not be attributed to an earlier developer without examining subsequent modification and the conditions of use.

Dataset preparation presents a different set of acts. A source custodian may have recorded a fact incorrectly; an annotator may have assigned a misleading label; an integrator may have combined two people into one record. Each event can contribute to an adverse result, but each has a different evidential location. The person able to correct the source may not possess the derived dataset, while the supplier maintaining the derived dataset may lack authority to amend the official register. Investigation must distinguish responsibility for the quality of the information from responsibility for deciding what follows from it.

Procurement and configuration connect these earlier choices to a particular institution. The procuring body determines what problem the acquired service is supposed to address and what evidence it can obtain when something goes wrong. The configuring actor may decide how missing values are treated, which threshold triggers intervention, and what an official sees. These choices should be reconstructed from the implemented configuration rather than the tender alone. A promised exception route that never appeared in the operational interface is not evidence that the official or affected person could use it.

Deployment and use concern the conversion of output into action. An official may accept a recommendation, seek further information, or decline to rely on it. A professional may incorporate a generated proposition into advice or a pleading. The record should show what was available to that person, what was actually considered, and the reason for the action. A later statement that “a human reviewed the case” is incomplete unless it identifies the review performed and the information on which it rested.

Four routes that must remain distinct

Administrative accountability concerns the lawful exercise and review of an institutional function. Its immediate objective may be a lawful procedure or reconsideration of an adverse act, rather than compensation from a developer. Professional discipline concerns the conduct of a regulated person under applicable professional rules. Contractual allocation concerns the obligations assumed by parties to an agreement. Civil liability concerns a legally recognised claim for loss or other injury against a defendant. These are working distinctions for organizing an investigation; the applicable instruments determine the jurisdiction, elements, defences, and remedies of each route.

The same incident can enter more than one route without producing the same conclusion in each. A hypothetical legal adviser may fail to verify a generated citation, while the supplier has complied with a narrowly defined contract. The professional issue is not resolved by contractual compliance. Alternatively, a supplier may breach an evidence-retention promise although the final decision is independently supportable. That breach does not automatically establish that the claimant lost the underlying case because of the AI tool. The investigation must resist both complete merger and complete separation of the routes.

South African data-protection law supplies a specific statutory illustration. Section 99 of the Protection of Personal Information Act 4 of 2013 permits the specified civil proceedings against a responsible party for a breach referred to in section 73, whether or not intent or negligence is present. The provision also identifies defences and permits compensation for patrimonial and non-patrimonial loss resulting from the breach. It does not impose liability on every participant merely because that participant contributed technology (POPIA, 2013, s. 99). The practical lesson is to identify the statutory defendant, qualifying breach, causal connection, and available defences instead of replacing them with a general claim of “strict AI liability.”

Brazil's General Personal Data Protection Law similarly requires attention to differentiated roles. Its article 42 addresses damage caused through processing contrary to data-protection legislation. It specifies circumstances in which operators and directly involved controllers are jointly responsible, while article 43 identifies grounds for exclusion. Article 42 also provides for recourse between responsible parties (LGPD, 2018, arts. 42-43). The analytical consequence is that a supply-chain map must be read together with the applicable legal allocation; commercial labels alone do not settle who can be sued or on what basis.

Causation and the difference between a defective output and compensable loss

An investigation should reconstruct the sequence from the disputed feature to the alleged injury. Consider a hypothetical benefits application linked to another person's record. A matching service returns a suspected duplicate. A local rule converts that flag into suspension. An officer approves the suspension without examining an identity document. The claimant incurs an additional expense while payment is interrupted. The matching error, the rule attaching consequences to it, the officer's action, and the expense are different propositions. Evidence supporting one does not automatically prove all the others.

Counterfactual questions can organize this inquiry without replacing the jurisdiction's causation test. Would the suspension have occurred with the correct record? Would an effective exception route have prevented it? Would the same expense have arisen for an independent reason? The answers require a comparison tied to the actual legal and operational conditions. Re-running the case through a newer model is not a complete counterfactual because the sources, configuration, and decision procedure may have changed. A reconstruction should identify each difference and explain which inference the experiment can support.

Causal uncertainty also needs to be localized. The evidence may establish that a particular record caused a suspension while leaving uncertainty about why the record was wrong. That uncertainty may matter to a damages claim against the data preparer but not to an immediate request for reconsideration based on corrected evidence. Conversely, a general error rate may support further investigation without establishing that the claimant's result was erroneous. The appropriate response is to narrow the claim and identify missing evidence, not to treat either individual experience or aggregate performance as conclusive for every issue.

An omission requires its own account. The investigator should identify the act allegedly required, the actor's opportunity to perform it, the source of the duty, and how performance could have affected the result. Failure to monitor, preserve evidence, provide a correction route, or communicate a known limitation can involve different duties. A statement that the harm “could have been prevented” does not establish that this defendant was legally obliged to prevent it in the manner proposed.

Human intervention, contractual allocation, and immunities

Retaining a human participant cannot be treated analytically as a complete supplier defence. A reviewer may have been given an interface that concealed uncertainty or omitted decisive source material. Equally, a reviewer may have disregarded a clear limitation and used the tool for an unauthorised purpose. The evidence should test both possibilities. Responsibility does not follow automatically from the last signature, but the signature also cannot be ignored when it records an exercise of legally assigned judgment.

Contracts can allocate verification, support, correction, insurance, and financial exposure between their parties. Their effect on an affected person outside the agreement is a separate legal question. Brazil's Consumer Defence Code illustrates the importance of statutory limits: article 51 addresses clauses that exclude or attenuate supplier liability for defects, subject to its qualification for justified limits involving a legal-person consumer. Whether the relationship falls within consumer law must first be established; the provision should not be generalized to every public-sector or business-to-business AI contract (Consumer Defence Code, 1990, arts. 2-3, 51(I)).

Immunity likewise requires an identified source and scope. Kenya's Constitution protects a member of the judiciary from an action or suit concerning an act or omission performed in good faith in the lawful performance of a judicial function (Constitution of Kenya, 2010, art. 160(5)). That wording cannot simply be extended to a supplier, every administrative employee, or every activity associated with a court. More generally, the inquiry must distinguish personal immunity, the validity of an act, review of a decision, and the powers of a disciplinary institution. A rule affecting one question does not answer the others by implication.

Lifecycle responsibility analysis can encourage indiscriminate claims if it treats every participant as a potential defendant without a defined legal basis. A properly completed matrix does the opposite. It requires a defined act, relevant rule, supporting evidence, and competent route before a participant is treated as a potential defendant. Some actors belong in the evidence map because they hold records, not because a cause of action against them has been established. Others may be necessary to implement a remedy even though financial responsibility remains disputed.

A false match across two legal relationships

Consider a constructed case, designated Case M. A matching service links an adult applicant to another person’s record. The institution treats the link as proof that the applicant is already receiving a benefit and suspends access. The applicant produces an identity document that contradicts the linkage, incurs an additional travel expense, and asks for restoration. The system’s stored output, the officer’s approval and the receipt for travel are available; the supplier has not yet explained the transformation that produced the link. No part of this narrative describes an observed deployment. Its purpose is to expose the additional premises required by different legal routes.

In the South African variant, stipulate that the suspension is administrative action within PAJA and that the institution is a responsible party for processing within POPIA. These are explicit jurisdictional assumptions, not conclusions that all public benefits or all court activities automatically fall within either Act. The claim concerns three objects: the accuracy of the personal record, the legality of suspension, and loss caused by the processing breach. A correction request may succeed without deciding the damages claim, while an administrative challenge may succeed because the institu-

tion used an irrelevant consideration even if the origin of the false link remains disputed.

For the processing claim, the relevant duty is concrete. POPIA section 16 requires reasonably practicable measures directed to information quality in light of its purpose. Sections 24 and 73 connect correction and qualifying interference to the enforcement framework. In Case M, the claimant would identify the erroneous linkage, its use and the verification opportunity that was available. A wrong output alone does not prove that the reasonable-steps duty was breached: the records must show what checking was practicable and why the safeguards used were inadequate for a decision that could interrupt support. The stronger allegation arises when contradictory identity evidence was available but never examined (POPIA, ss. 16, 24, 73).

The statutory damages route is against the responsible party under section 99. It does not require proof of intent or negligence, but it still requires the qualifying breach and loss resulting from it. The institution's role must be established from its decisions about the purpose and means of processing; the vendor's product name is not a substitute for that inquiry. The travel receipt supports an amount, but the claimant must also connect the journey to the erroneous suspension. If the journey would have occurred for an independent verification requirement, the whole expenditure cannot simply be attributed to the false match. Any non-patrimonial claim likewise needs an account of the injury rather than a formula equating every inaccurate record with a fixed award.

Section 99(2) expressly permits defences concerning vis major, the claimant's consent or fault, reasonable practicability of compliance, and a section 37 exemption. Their availability must be tested against the act alleged. Consent to necessary benefit administration is not analytically identical to consent to a false linkage; an institution invoking consent must explain its scope. Outsourcing the matching function is not itself one of the enumerated defences. Conversely, a claimant's provision of inconsistent identifiers may bear on a pleaded defence and on causation. These are applications proposed for the constructed facts, not predictions of a decided case. Section 99 places the damages action in a court having jurisdiction; it does not make the Information Regulator an automatic compensation tribunal (POPIA, s. 99(1)-(3)).

The administrative route asks a different question. Failure to consider the contrary identity document could engage PAJA section 6(2)(e)(iii); unfair exclusion from participation could engage section 6(2)(c); and a suspension disconnected from the relevant information could engage section 6(2)(f)(ii). The pleaded ground should identify the legal defect in the actual act rather than allege that the institution used an algorithm. The proposed relief would ordinarily seek lawful reconsideration with the corrected record, accompanied by any justified interim protection. Substitution and compensation under section 8(1)(c)(ii) are exceptional remedies, not the automatic consequence of showing an error (PAJA, ss. 6, 8).

The temporal analysis must be separated by claim. PAJA's section 7 review period and exhaustion conditions are examined in section 4.4. A money claim requires a separate prescription analysis: the Prescription Act's residual period is three years, subject to another statutory rule, and commencement turns on a due debt and the knowledge provisions in section 12. Delay, acknowledgment and judicial interruption have their own conditions. For a damages claim falling within the Institution of Legal Proceedings against Certain Organs of State Act, written notice under section 3 ordinarily must be served within six months after the debt became due; the Act also provides consent and condonation routes. Neither period should be inferred from the date of a technical audit. The defendant, debt, knowledge and applicable notice provisions must be pleaded from the actual claim (Prescription Act, 1969, ss. 11(d), 12-16; Institution of Legal Proceedings Act, 2002, ss. 1, 3-4).

The Brazilian variant preserves the mechanism but changes the legal relationship. Stipulate a natural person who obtains a paid private service as its final recipient, a supplier within the Consumer Defence Code, and consequential harm from a false-match suspension. This is a service-defect claim, not an assertion that a public-benefit recipient is necessarily a consumer. Article 14 makes supplier responsibility for damage caused by defective service independent of fault; the expected safety of the service remains the relevant inquiry. A matching tool used only to prompt a reversible identity check presents a different service risk from one that closes access despite contradictory documents. The claimant must connect the challenged service, defect and damage, while the supplier may prove absence of defect or exclusive fault of the consumer or a third party under article 14(3) (Consumer Defence Code, arts. 2-3, 14).

The LGPD supplies a concurrent processing inquiry without making every subcontractor automatically liable. Article 42 addresses damage caused through processing in breach of data-protection legislation and specifies when operators and directly involved controllers share responsibility. Article 43 provides exclusions where the actor did not carry out the attributed processing, there was no violation, or damage resulted exclusively from the data subject's or a third party's fault. Article 45 preserves the applicable consumer-liability rules. Accordingly, the pleadings can distinguish the service supplier's consumer obligation from the identified controller's processing conduct. They should not assume that contractual compliance between suppliers extinguishes either duty, or that the same compensable loss can be recovered twice (LGPD, arts. 42-45).

For the stipulated article 14 damages claim, article 27 supplies a five-year prescription period running from knowledge of the damage and its authorship. That rule must not be exported to every LGPD, contractual or service-quality dispute: article 26 deals separately with the time limits for specified defect complaints. Article 101(I) permits a supplier-liability action at the claimant's domicile; the competent court and suitable procedure still depend on the parties, subject matter and claim. The worked route assumes an ordinary state-court consumer action involving a private supplier, rather

than a claim requiring federal jurisdiction. If immediate reinstatement is sought, CPC article 300 directs attention to the probability of the right and danger of harm or risk to effective proceedings, together with its safeguards, not to a completed forensic explanation (Consumer Defence Code, arts. 26-27, 101(I); Code of Civil Procedure, art. 300).

The comparison produces a limited conclusion. The same false match can be examined through public-law validity, processing duties or consumer-service responsibility, but the elements and procedural clocks do not travel together. In both variants, the institution's decision record and the claimant's account can establish a sufficiently concrete dispute before the supplier's internal mechanism is known. They do not dispense with proof of the specific duty, causal connection and remedy. The responsibility matrix therefore records a legally formulated claim, including a reason to exclude an actor, rather than merely drawing an extensive technical supply chain.

Responsibility matrix

Table 4.1 is an investigative instrument, not a finding of liability. A legal reviewer should replace the potential-duty descriptions with the actual provision, professional rule, or contractual term. The case record should distinguish allegations, established facts, disputed facts, and matters outside the available evidence. Where an actor occupies several roles, each role should be assessed separately rather than merged under a product or corporate name.

Table 4.1. Responsibility across an AI-assisted decision process

Actor and conduct to examine	Potential duty requiring legal identification	Relevant evidence	Possible enforcement or corrective route
Developer or model supplier: design, claims, warnings, updates	Agreed performance or disclosure obligation; applicable statutory or civil duty	Specifications, validation scope, release history, representations, known limitations	Contract enforcement; applicable civil or regulatory proceedings; technical correction
Source custodian or dataset preparer: entry, labelling, linkage, reuse	Applicable accuracy, lawful-processing, custody, or contractual requirement	Source records, provenance, annotation instructions, transformation history, correction notices	Source correction; data-protection process; contractual or civil claim where established

Table 4.1. Responsibility across an AI-assisted decision process

Actor and conduct to examine	Potential duty requiring legal identification	Relevant evidence	Possible enforcement or corrective route
Procuring and deploying institution: selection, authorisation, supervision	Mandate, procedural duties, agreed supervision and evidence-access obligations	Approvals, tender, contract, risk review, staffing, complaints, audit access	Administrative review; institutional investigation; procurement or contractual enforcement
Integrator or configuration owner: thresholds, interfaces, exception rules	Contractual conformity and any applicable legal duty concerning implementation	Configuration versions, change approvals, missing-data rules, interface captures	Reconfiguration and reprocessing; contract claim; other liability if its elements are met
Official or professional user: reliance, verification, reasons	Assigned decision function; applicable competence, candour, supervision, or procedural duty	Material actually reviewed, source checks, overrides, reasons, training and workload records	Reconsideration or appeal; professional or administrative process; personal claim subject to applicable limits
Host or technical operator: preservation, access, security, interruption	Applicable processing, security, retention, disclosure, or continuity obligation	Access logs, retention settings, incident reports, service records, preservation requests	Regulatory or contractual enforcement; restoration of records and service
Reviewing body or remedial owner: handling and implementing correction	The body's own mandate and any enforceable review or implementation duty	Complaint chronology, interim decisions, orders, restoration evidence, unresolved referrals	Escalation within jurisdiction; review or enforcement of an existing obligation

Completing the matrix should produce a short theory of the case for each proposed route: the alleged conduct, the rule engaged, the evidence available, the injury or procedural defect, the respondent, and the requested outcome. It should also identify who can act immediately. A dispute between suppliers over reimbursement need not be treated as a reason to postpone every operational correction. Whether a particular interim or final measure is legally available remains a question for the competent institution, examined in section 4.4.

4.2 Notice, Reasons, and Access to Decision Records

What an affected person needs to know

A person cannot formulate an effective challenge from the statement that an application was "processed by AI." That description identifies a possible technique, not the act being challenged, its legal basis, or the information that mattered. A useful notice must begin with the outcome: what has been refused, suspended, restricted, assessed, or required; when the action takes effect; and which institution is responsible. Only then can the role of the system be placed within the decision rather than presented as an explanation in its own right.

The necessary information depends on the ground of challenge. A person alleging mistaken identity needs the relevant identifying information and the basis of the match. A person disputing a deadline needs the rule, triggering event, recorded date, and calculation. A person alleging that an irrelevant criterion determined the outcome needs the criterion actually applied and the evidence linking it to the action. A claim that the institution failed to consider an exception requires information about both the exception and the treatment of the person's explanation. These are decision-specific inquiries, not demands for undifferentiated technological transparency.

System explanation and legally adequate reasons

An explanation of system behaviour might show how an output changes when an input changes. A procedural explanation might show how the output enters a queue or recommendation. Legal reasons must address a different question: why the responsible institution reached this outcome under the applicable criteria and evidence. The three accounts can assist one another, but none should be treated as a substitute for the others. A feature-importance display does not establish that the highlighted feature was legally relevant. A technically faithful account of a classification does not explain why classification justified exclusion.

This distinction also limits the use of generated explanations. A fluent narrative written after the event may describe a plausible basis for the result without recording the reasoning actually used. The investigator should compare the narrative with contemporaneous inputs, instructions, output, review, and reasons. Where the institution has reconsidered the case and adopted a new reason, the record should say so. It should not present a later reconstruction as an unchanged account of the original decision.

South Africa's Promotion of Administrative Justice Act illustrates the connection between procedure and reasons within its defined field. Section 3 addresses procedural fairness, including notice and opportunities to make representations, subject to its qualifications. Section 5 provides a route to request written reasons and identifies consequences and permitted departures where adequate reasons are not supplied (PAJA,

2000, ss. 3, 5). The relevant analytical distinction is between understanding an administrative act and understanding its software; the Act's applicability must still be established for the act in question.

In the hypothetical duplicate-record suspension, a sufficient operational explanation might state that a matching service returned two linked entries. The legally relevant account must go further: which eligibility or verification requirement was involved, whether a duplicate justified suspension or only further inquiry, what contrary evidence was considered, and who decided the matter. This structure allows the affected person to challenge the inference from record to consequence without having to prove that the entire matching system is defective.

Establishing the material AI role

Materiality should be assessed by tracing what the AI contribution could change in the particular process. It may select records for consideration, summarise evidence, recommend a result, generate reasons, or execute an action. A source-selection tool can matter even when it does not write the final decision. Conversely, routine formatting does not necessarily influence the substance. The proposed disclosure protocol therefore records the contribution's function and effect rather than using a binary statement that AI was or was not present.

A notice should not obscure differences between a generated output and the institution's adoption of it. The system might have identified a risk while the official decided to request information, or the system might have closed the application automatically. These are different acts with different opportunities for intervention. Where a person reviewed the output, the record should describe the scope of that review. The phrase “human oversight” is not informative unless it identifies what the reviewer could inspect and change.

The transparency analysis in T-323/24 emphasises explaining judicial AI use in a way that enables understanding, including the tool's role and limitations. That reasoning belongs to the Colombian case, not to a universal disclosure code (Corte Constitucional de Colombia, 2024, para. 371). Its comparative value is the question it prompts: can a party identify the contribution that may require an answer, rather than being asked to trust an unexplained assurance of responsible use?

Decision records, source code, and graduated requests

Decision-specific disclosure should normally begin with the material necessary to reconstruct the outcome: the operative act, relevant inputs, source dates, criteria, output, version, human intervention, and notice. The request can then be expanded where those records leave a material issue unresolved. A disputed matching threshold may require configuration information. A suspected transformation error may require the relevant mapping or code. A question about a model's performance may require validation evidence and a suitably bounded test. This is a sequence for framing requests, not a rule restricting a court's or regulator's powers.

Source code can be necessary to an individual challenge when the disputed defect concerns the implemented rule and cannot otherwise be examined adequately. Equally, source code alone may be insufficient. It may omit the deployed configuration, model parameters, retrieved documents, external service response, or intervention that produced the actual output. The request should therefore identify the missing evidential link. Asking for “the algorithm” without identifying the issue can produce a large technical disclosure that still fails to explain the disputed act.

Brazil's LGPD offers a specific illustration of qualified transparency. Article 20 addresses review of decisions based solely on automated processing that affect the data subject's interests. Its first paragraph requires clear and adequate information about decision criteria and procedures while protecting commercial and industrial secrecy; its second permits the national authority to audit for discriminatory aspects when information is withheld on those grounds. The operative text does not itself promise review by a natural person (LGPD, 2018, art. 20). This is a defined statutory arrangement, not a general entitlement to public source-code disclosure.

South Africa's POPIA is differently framed. Section 71 addresses qualifying solely automated profiling decisions and identifies exceptions. For the protective-measures exception connected with a contract under section 71(2)(a)(ii), subsection (3) requires an opportunity to make representations and sufficient information about the underlying logic to do so (POPIA, 2013, s. 71). The precise scope matters: that subsection should not be detached from its statutory setting and presented as a universal explanation right covering every AI-assisted act.

Confidentiality, security, and protected information

A trade-secret objection should identify the material claimed to be protected and the legal ground for protection. A security objection should explain the risk associated with the particular disclosure. A privacy objection should distinguish the requester's information from information about others. These distinctions permit an adjudicator or other competent body to consider a targeted response instead of choosing only between unlimited publication and complete secrecy. Whether that body can order a particular form of access depends on the applicable procedure.

Possible arrangements for consideration include redaction of unrelated personal information, disclosure to an authorised representative, inspection by an independent expert, a protected testing environment, or a summary accompanied by review of the withheld material. These are proposed procedural options requiring local legal authority. They should be tested for effectiveness: an expert who can inspect but cannot communicate any meaningful conclusion to the affected person may not resolve the underlying participation problem.

Confidentiality should not be reduced to a contest between a claimant and a supplier. Other affected people may have legitimate interests in records that reveal sensitive circumstances or confidential communications. A broad request for all training records can expose people who are not parties to the dispute. The request should explain the relevance of each category, identify safer alternatives where adequate, and preserve the ability to seek further access if the alternatives cannot answer the issue. Minimization and effective challenge can be considered together rather than assumed to be incompatible.

Accessible notice and reliable procedural information

Notice must be assessed from the recipient's position. The institution may have generated a letter without establishing that the person could obtain, understand, or act on it. The proposed protocol records the chosen channel, delivery evidence, language, accessibility needs, and available assistance. It also distinguishes the date of the act from the date on which the recipient was notified. Neither should be silently substituted for the other when calculating a procedural period.

Review instructions deserve the same verification as the substantive decision. An incorrect forum or deadline can frustrate an otherwise useful explanation. A model-generated statement about appeal rights should therefore be checked against the applicable procedure before release. Where different periods govern internal review, information access, judicial proceedings, or a damages claim, they should be recorded separately. The protocol does not assume that requesting records, lodging a complaint, or negotiating with a supplier suspends another period.

Detailed notices and disclosure can impose substantial burdens on institutions, particularly for large volumes of routine decisions. Proportionate design can address this without reducing every reason to a generic code. Standard wording can explain a common procedure; case-specific fields must identify the actual information, criterion, effect, and review route. The institution should avoid generating an apparently individualized account from fields that do not accurately record the decision. A shorter truthful notice is more useful than a lengthy explanation whose relationship to the event cannot be demonstrated.

Model notice and record-request template

The following models are not prescribed statutory forms. They are intended for adaptation by an authorised institution, legal-aid provider, lawyer, or affected person after checking the jurisdiction, confidentiality requirements, competent recipient, and procedural periods. Bracketed entries are fields to complete, not conclusions about existing rights. Neither model substitutes for filing an appeal or seeking urgent relief where a separate application is necessary.

Model 4.1. Notice of a decision involving an AI system

Decision and effect. On [date], [institution and competent office] decided [specific outcome] in matter [reference]. The decision takes effect on [date] and affects [payment, application, account, obligation, or other interest]. It is [provisional/final/subject to the stated further step].

Authority and reasons. The relevant authority and criteria are [instrument and precise provision]. We relied on [material facts and source dates]. We accepted or rejected [important contrary evidence or explanation] because [case-specific reason]. The conclusion follows from those criteria and facts as follows: [reasoned connection].

System contribution. [System and version, where known] performed [search, matching, classification, prediction, generation, or other task]. Its relevant output was [output and any material limitation]. [State whether action was automatic or human-reviewed; identify the responsible office, material examined, action taken, and authority for that procedure.]

Records and disputed information. The attached or accessible records are [list]. Requests for further material or correction should be directed to [contact and applicable procedure]. Where material is withheld, [specific legal ground and route to challenge the withholding] applies.

Review, urgency, and assistance. A challenge may be brought to [competent body] by [verified date, triggering event, and procedure]. Other relevant periods are [separate entries]. [Procedure] is available to request urgent consideration or interim protection where its legal conditions are met. Assistance and non-digital or accessible channels are [details]. The effect of a record request on any review deadline is [verified position, with the applicable rule or order].

Model 4.2. Request for decision records and preservation

Identification and purpose. I request records concerning [decision/reference/date] by [institution]. I seek to understand and, where appropriate, challenge [specific issue]. My authority to request the material is [applicable access, procedural, representative, or other basis; supporting authorisation where necessary].

Requested material. Please provide the operative decision and reasons; the relevant input records and their sources and dates; the criteria and applicable rule or configuration; the material system output; the deployed version identifier; the record of human review, override, or escalation; and the notice and delivery record. The connection between each requested category and the disputed issue is [explanation].

Further technical information. The unresolved technical issue is [identified issue]. Please provide [specified configuration, validation evidence, relevant transformation, or other material] needed to examine it. Where a source-code or protected-material request is necessary, its scope and proposed safeguards are [details].

Preservation and missing records. Please preserve relevant existing records, versions, and correspondence through a lawful, proportionate hold while this matter is considered. If a requested item is unavailable, identify whether it was never created, lost, deleted, held elsewhere, or withheld, together with the available explanation and relevant custodian. This request does not authorise unlawful access or unrestricted retention.

Response and procedure. Please identify the applicable response period, any lawful charge, available assistance, and the route to contest refusal or incomplete disclosure. For withheld material, provide the legal basis and consider severance or protected access where permitted. I separately request [urgent action, where appropriate] and do not assume that this records request replaces or extends any other procedural step.

The issuer or adviser should retain the completed template, the legal checks supporting it, and evidence of transmission. A reviewer should then ask whether the recipient can identify a concrete ground of challenge and reach the appropriate body. The quality of notice is measured here by what it enables a person to do, not by the volume of technical terminology it contains.

A completed request in the constructed case

The following specimen completes the parts of a notice that can be supplied without inventing an empowering statute or a review forum. It is a constructed request for reasons and preservation in the South African variant of Case M, not a prescribed form or a notice issued by an existing authority. Assume that the applicant first learned of the suspension on 25 September 2026, received no adequate reasons, and sends this request on 2 October 2026. The dates, initials and reference are fictitious. Unlike an institutional notice defending the suspension, the applicant's request need not invent the provision that the administrator failed to identify.

To the administrator responsible for Case M001. I request written reasons under section 5 of the Promotion of Administrative Justice Act for the suspension communicated to me on 25 September 2026. I became aware of the suspension on that date. The message stated only that a duplicate record had been identified. It did not identify the eligibility provision, the evidence establishing a duplicate entitlement, the official who

decided the matter, or a route for challenging the decision. I request reasons addressing those matters and explaining why my identity document did not resolve the mismatch.

My dispute concerns the linkage of my application to a record belonging to another person. Please identify the inputs and source dates used for the match, the material matching result, the rule connecting that result to suspension, and the record of the official's review. Please preserve the existing version and decision records relevant to this event. I seek only information necessary to understand my own matter; unrelated personal information about the other person may require protected treatment. If you consider a category unavailable or lawfully withheld, please identify the category, custodian and reason, so that I can use the appropriate access procedure.

I also request urgent consideration of whether the suspension can lawfully be paused while the identity discrepancy is examined. The attached account explains the effect of interruption and the alternative identity evidence available. This request does not assume that section 5 itself suspends the decision or grants an interim remedy. Please identify any statutory internal remedy, its recipient and period, and confirm which office can decide the urgent request. I will retain the transmission acknowledgment and the attachments with the original suspension message. Signed: Applicant M, 2 October 2026; reply through the assisted-service channel used for this submission.

The accompanying adviser's record would note that a request made seven days after awareness is within section 5(1)'s ordinary 90-day request period. Section 5(2) gives the administrator 90 days after receipt to provide adequate written reasons, subject to the Act's qualified departures; it is not a promise of immediate restoration. Review under section 7 must be assessed separately, including applicable internal remedies and unreasonable delay. A reasons request is therefore filed promptly alongside, rather than instead of, the work needed to protect the review route. This specimen demonstrates a completed, issue-based request while preserving the distinction between a verified statutory period and a deadline that cannot yet be calculated from the available facts (PAJA, ss. 5, 7, 9).

4.3 AI-Related Evidence and Standards of Proof

Three evidential objects

AI-related evidence should first be separated into material produced by a system, material altered through a system, and records about how a system operated. The first category includes a generated summary, classification, recommendation, or synthetic recording. The second includes an enhanced image, translated statement, edited recording, or automatically corrected transcript. The third includes input records, versions, settings, retrieval results, logs, human interventions, and communications concerning deployment. The distinction identifies what must be proved and prevents a convincing output from being treated as evidence of its own reliability.

The intended evidential use matters as much as the category. A generated letter may be authentic evidence of what an institution communicated even though the legal propositions in it are false. A transcript may show what a system produced without establishing exactly what a witness said. A risk score may establish that the institution received a recommendation without proving the underlying allegation. The proponent should identify the proposition for which the material is offered. The opposing party can then contest that proposition rather than engage in an abstract debate about whether AI can ever produce useful evidence.

Authentication, admissibility, reliability, and weight

Authentication asks whether the item is what it is claimed to be. Admissibility concerns whether the applicable legal procedure permits its use. Reliability concerns the dependability of the item or method for the relevant inference. Weight concerns the significance assigned to it alongside the rest of the record. These working distinctions prevent one successful check from answering questions it cannot resolve. An accurately preserved output can remain substantively wrong; a useful technical test can remain insufficiently connected to the disputed event.

In Shafi Mohammed Sarki v The State, the Sindh High Court examined proof and preservation of modern-device evidence in a narcotics appeal. Paragraph 13 reproduces the Supreme Court's statement in Mian Khalid Pervaiz that an automated record, when denied, must still be proved according to law; paragraphs 12-13 relate that inquiry to the Qanun-e-Shahadat framework and the Electronic Transactions Ordinance. The source used here is the High Court's own judgment, not an independently examined full report of the earlier Supreme Court case. The appeal concerns an evidential setting that can illuminate AI-related questions but does not decide admissibility of synthetic material (Sindh High Court, 2025, paras. 10-13).

The practical implication is to prepare both a production account and an inference account. The production account identifies where the record came from, how it was acquired, and what happened to it. The inference account explains why the record sup-

ports the factual proposition asserted. An expert can assist with either, but the report should not move silently from identifying a system output to validating the allegation that the output expresses.

Domestic foundations for a proposed inference

The inquiry in Case M can be organized around four proposed inferences. An institutional notice may prove that suspension was communicated; a matching log may prove that a service returned a flag; the source record may support or contradict the claimed identity; and an expense record may support a loss claim. One electronic file may be offered for more than one inference, but its admission for one purpose need not validate the others. A generated account stating that two applications concern the same person cannot obtain the status of reliable identity evidence merely because its export is authentic.

India's Bharatiya Sakshya Adhiniyam entered into force on 1 July 2024. Sections 61-63 govern electronic-record treatment, while section 170 saves the earlier regime for the pending matters it describes. A party must identify that temporal basis before choosing its production method. Where section 63 applies, the conditions concern regular use, the information supplied, proper operation and the relationship between the output and the information processed; subsection (4) and the Schedule specify the certificate mechanism. The Schedule has party and expert parts. A generic vendor assurance does not answer those particulars (Bharatiya Sakshya Adhiniyam, ss. 61-63, 170, Schedule; Ministry of Home Affairs, 2024, S.O. 849(E)).

This production framework has a limited application to a generated summary. The certificate can identify the device or service record and the conditions under which the export was produced. It cannot retrospectively make an invented source statement true. Section 39 separately addresses expert opinion, including electronic-evidence expertise, and sections 104-106 address burdens attached to asserted facts. The proposed litigation record should consequently separate the custodian's account of extraction from any expert's opinion about matching accuracy and from the claimant's evidence about identity. These are applications of distinct statutory questions; the chapter does not claim an Indian judgment has already approved this exact protocol for generative outputs.

South African hearsay law supplies another necessary distinction. Section 3 of the Law of Evidence Amendment Act defines hearsay by dependence on the credibility of someone other than the witness and provides routes through agreement, testimony by that person, or the court's interests-of-justice determination. Its factors include the proceeding, evidential purpose, probative value, explanation for absence and prejudice; overcoming hearsay does not cure a different inadmissibility ground. A summary reproducing an absent officer's account therefore calls for an analysis of the human source, not simply certification of the summary file. An inference generated independently by software presents a different foundation question; calling all machine materi-

al hearsay obscures the statutory definition (Law of Evidence Amendment Act, 1988, s. 3).

The practical dispute may therefore concern the missing witness, the process that transformed the source, or both. Suppose the log records a clerk's statement that the applicant admitted a duplicate entitlement, but the applicant denies making it. An expert who can authenticate the database entry has not resolved the clerk's credibility. Suppose instead the system inferred identity from similar names without a human statement. The issue shifts toward the method, inputs and error conditions. The advocate should identify the alternative route by which the same fact could be proved, including the source document or appropriate witness, before requesting a costly examination of the complete model.

The Pakistani illustration also contains a tension worth confronting. Sarki paragraph 10 strongly endorses video recording of narcotics recovery and describes circumstances in which prosecution evidence would require an answer from the defence. Paragraphs 12-13 nevertheless demand proper preservation and proof, and paragraphs 17-19 apply the criminal burden and acquit. The favourable description of recording technology cannot be detached from those qualifications and converted into a presumption that a recorded event is genuine or that an accused must disprove an unverified synthetic recording. This is the chapter's critical reading of the judgment's internal sequence, not a holding on a later AI dispute (Sindh High Court, 2025, paras. 10-19).

The distinction between an evidential response and the legal burden must remain visible. Credible prosecution material may call for an explanation without turning a detector's confidence into proof beyond reasonable doubt. A respondent in a civil case may possess easier access to a particular technical fact without being required to disprove every allegation about a model. In Brazil, CPC article 373 permits a reasoned, qualified reallocation tied to difficulty and access, while LGPD article 42(2) supplies a specific judicial burden-reversal provision in its own field. These are powers to be argued and exercised through procedure, with notice and a real opportunity to respond. They are not automatic effects of the word opacity (Code of Civil Procedure, art. 373; LGPD, art. 42(2)).

For an expert examining Case M, the material question is whether the retained input could have produced the disputed link under the deployed configuration, and what competing explanations remain. The report should expose the reference data, validation population, threshold, missing inputs and selection of test examples. If the historical service cannot be run, the expert can compare documented rules and outputs while stating that causal reproduction was unavailable. A proposition about a later model should not be offered as a historical fact. Equally, failure to reproduce a stochastic output does not demonstrate fabrication of the preserved record. The proposed method keeps the limits of technical inference available for cross-examination and for the court's separate judgment about weight.

Provenance and the reconstruction of an event

The W3C PROV framework organizes provenance through entities, activities, and agents, together with relationships between them. It provides a technical vocabulary for recording how information was produced and transformed, rather than a legal certification of its truth (World Wide Web Consortium, 2013, §§1-2). That distinction is useful for an AI dispute: a claimant needs an intelligible sequence of records and actions, not merely a collection of files carrying the same case number.

A reconstruction should identify the source item, the acquired copy, each material transformation, the responsible actor or process, and the output supplied to the decision-maker. Dates and time zones should be recorded where relevant, together with uncertainty about unsynchronized clocks. The investigator should distinguish a log timestamp from a verified event time. Where several services communicate, the record should connect request and response identifiers instead of assuming that entries with similar times describe the same transaction.

The preservation protocol proposed here uses a protected original or acquired reference copy and separately identified working copies. Hash values can assist comparison of the copies, but a matching value does not prove that the original account was truthful or that the material had not already been altered before acquisition. The record should state what was hashed, when, by whom, and with which method. It should also document any limits on acquisition, such as receiving an exported file without access to the original device or service.

Preservation must not itself become an unauthorised investigation. A request to retain relevant records does not permit bypassing access controls, copying unrelated client files, or obtaining credentials through deception. The authorised custodian should specify the material held, access restrictions, preservation period, and review of the hold. Where legal duties require retention of a disputed record, correction may need to preserve the historical version while identifying the current account. Erasing the history can make it harder to establish both the error and the action taken to repair it.

Synthetic recordings, altered material, and detector limitations

Synthetic origin is not the same as falsity of every proposition conveyed, and non-synthetic origin is not proof of an accurate account. A demonstrative reconstruction openly identified as such has a different evidential purpose from a recording presented as an unaltered event. The party relying on the material should identify whether it is offered as a record, an illustration, an expert reconstruction, or something else. The receiving body should examine the claim actually made about it.

NIST's 2024 technical overview distinguishes provenance, watermarking, detection, and labelling approaches. It explains that detectors can be wrong, their uncertainty estimates may be miscalibrated, and even signed metadata can be incomplete or inaccurate. It also cautions against assuming that synthetic material is necessarily deceptive

while non-synthetic material is benign (Chandra et al., 2024, §§3.3.1.1, 3.3.1.3-3.3.1.4). These are technical cautions, not rules governing admissibility or the standard of proof in the focal jurisdictions.

A detector result should consequently be treated as a claim requiring its own foundation. What did the tool test? Which version and settings were used? What material was submitted? What does its score represent? What evidence supports its suitability for that kind of text, audio, image, or video? An unsupported label of "AI-generated" should not displace examination of source records, witnesses, contextual consistency, and the processing history. The absence of such a label is equally insufficient to establish that no alteration occurred.

Enhancement and translation require a parallel account. An investigator who denoises an audio file or improves an image should retain the input and identify the transformation. The enhanced output may make a feature easier to inspect, but the report must distinguish information recovered from the source from information introduced or inferred by the method. A disputed translation should similarly preserve the original, the automated rendering, and any human revision. The court or reviewer needs to know which version supports the proposed inference.

Fabricated citations and generated factual accounts

A fabricated citation presents two different problems. As a legal authority, it cannot support a proposition if the judgment does not exist or does not contain the attributed holding. As a record of professional conduct, the document containing it may be relevant to what was submitted, checked, or represented. The investigation should therefore preserve the original document and its transmission history while independently verifying the supposed authority. Replacing the citation in a revised filing should not obscure the sequence under review.

A generated summary of evidence requires an equally disciplined comparison. The relevant question is not whether the summary sounds plausible, but whether it accurately represents the underlying material for the legal task. The checking record should identify omitted qualifications, changed attribution, unresolved ambiguity, and any proposition for which no source exists. Where the original evidence is missing, a summary cannot be treated as a complete substitute simply because it was produced before the dispute arose.

An institutional explanation generated after a complaint is a further evidential object. It may assist understanding, but the investigator should establish whether it was derived from preserved records, a later interview, or a model's general account of likely system behaviour. The explanation's date and method should be explicit. A reconstruction that incorporates assumptions can still be useful if those assumptions are disclosed; it becomes misleading when it is presented as a contemporaneous record.

Versions, experiments, and human intervention

A system that is the subject of a complaint may change while the complaint is being reviewed. Reconstruction therefore needs to establish which models, service versions, configurations, relevant data snapshots, input prompts, retrieval sources and connected tools were used at the time. It should also establish what the user viewed, whether the output was edited and whether an automated recommendation was considered. A consistent product name does not establish that the deployed system or its configuration remained unchanged.

When attempting to recreate an issue (i.e., replicate the original event), care should be made to differentiate between replaying a historical request versus running a new experiment. Running the same request against a current production service may produce a different result and may not accurately represent the original incident. A saved output still represents the output provided at the time; simply having a different output at a later date does not independently confirm that the previous output is incorrect. The recreated report should detail which elements were kept the same throughout the recreation process, which elements were unattainable, and which conclusions can and cannot be supported based on the recreation.

Do not view human interactions as minor details. Official corrections, translator revisions, attorney selections amongst generative drafts, administrator changes to queues -- all of these actions can affect the functional impact. In addition to identifying those events and their sequence, the evidentiary record must provide sufficient context so that absent documentation does not equate with absence of action.

Missing records and the applicable burden of proof

The consequences of missing evidence depend on the applicable law, the relevant preservation or production duty, the circumstances of its absence and its importance to the dispute. Evidence may never have existed, may have been automatically deleted, may be held by another entity, or may be withheld on a claimed legal ground. Intentional destruction, negligent loss and lawful disposal require different analysis. The party seeking a procedural consequence should identify its legal basis and explain how the missing material affects the case. An unexplained absence can itself be relevant; technical opacity does not automatically reverse the burden of proof.

Brazil's Code of Civil Procedure supplies a structured approach to the burden of proof. Article 373 ordinarily assigns proof of facts constituting the claimed right to the claimant and proof of facts preventing, modifying or extinguishing that right to the defendant. In the specified circumstances, a reasoned judicial decision may redistribute the burden because proof is impossible or excessively difficult for one party, or easier for the other. The affected party must have an opportunity to discharge the reassigned burden, which must not itself become impossible or excessively difficult (Code of Civil

Procedure, 2015, art. 373). This is a general procedural rule, not an AI-specific exception.

In addition, the way a court addresses uncertainty must align with the nature of the proceeding. For example, the potential adverse effects associated with an evidentiary shortcoming related to a civil damages action may not be the same as those in a criminal prosecution, administrative review, or request for provisional relief. The Sindh High Court appeal discussed above resulted in acquittal because the prosecution evidence left reasonable doubt, including doubts about recovery and safe custody; it did not concern AI-generated evidence (Sindh High Court, 2025, paras. 9, 14-19).

Similarly, a model's numerical confidence rating should not be treated by a court as being equivalent to its own standard of proof. A score's meaning depends on the model and its validation, and it may be poorly calibrated; the legal burden of proof relates directly to the propositions that the applicable law requires each party to prove in support of their respective claims.

Finally, requiring an exhaustive historical record of a model's development may impose disproportionate litigation costs. As such, courts should allow for phased investigation based on discrete substantive issues. There may be a simple source-record correction that resolves all or most of the key issues without requiring full disclosure about the model's architecture. On the other hand, there may be disputes over systemic changes within the model where more comprehensive technical discovery would be warranted. The party seeking additional evidence should provide clear articulation of the disputed proposition, explain why the evidence is relevant and why a less burdensome alternative would be inadequate, subject to the applicable disclosure rules. Proportionality refers to how broad or narrow the investigation should be permitted - not whether or not a party has disguised their lack of knowledge / uncertainty as some form of certitude.

Evidence-preservation checklist

Table 4.2 should be completed by an authorised records custodian with legal and technical input. Each entry should identify the actual record and responsible person, not merely indicate that a policy exists. The checklist is a proposed preservation instrument; local admissibility rules, disclosure duties, legal holds, and confidentiality requirements remain controlling.

Table 4.2. Preservation requirements for an AI-related dispute

Record category	What to preserve or document	Integrity and scope check
Disputed act and communications	Decision, original notice, submission, delivery record, complaint and requested relief	Retain the version actually communicated and its date

Table 4.2. Preservation requirements for an AI-related dispute

Record category	What to preserve or document	Integrity and scope check
Source material	Relevant records, recordings, documents, provenance and source dates	Distinguish original, acquired copy, export and unavailable source
System event	Relevant input, output, request identifier, timestamp and interface presented to the user	Connect records to the specific event; record clock uncertainty
Deployed configuration	Model/service identifier, rules, thresholds, prompts, retrieval and tool settings where material	Identify known version limits and subsequent changes
Human intervention	Reviews, edits, overrides, translations, approvals and reasons	Separate contemporaneous records from later recollection
Transformations and testing	Working copies, methods, parameters, test data, results and unsuccessful tests	Preserve the baseline; identify assumptions and limitations
Custody and access	Acquiring person, authority, transfers, access history, hashes where used	Protect originals and explain what integrity checks establish
Unavailable or withheld material	Custodian, retention policy, deletion event, access objection or search undertaken	Distinguish non-creation, loss, destruction, non-possession and refusal
Continuing preservation	Scope of hold, responsible officer, security, review date and authorised disposal	Avoid unrestricted retention and preserve unresolved-dispute material lawfully

Expert-question schedule

The questions in Table 4.3 should be adapted to the expert's field and the disputed proposition. Technical expertise should not be used to imply expertise in every relevant legal, linguistic, statistical, or institutional issue. The instructing party should disclose material that could qualify the desired conclusion and record the expert's independence and access limitations.

Table 4.3. Questions for an expert examining AI-related evidence

Inquiry	Questions the report should answer
Scope and competence	What precise proposition are you examining? Which parts fall within your expertise, and which require another specialist?
Material received	What did you receive, from whom, under what conditions? Which originals or relevant versions were unavailable?
Provenance	Can the item's origin and material transformations be traced? Which links are documented and which are inferred?
Method and validation	Why is this method suitable for the disputed material? What tests, limitations, and alternative explanations were considered?
Synthetic-content assessment	What does any detector score mean? What supports its use for this material, and what uncertainty remains?
Historical reconstruction	Which model, data, settings and human interventions can be established? Is the exercise a replay or a new experiment?
Causal contribution	What evidence connects the identified feature to the output and subsequent act? What competing explanation remains plausible?
Missing evidence	How does each missing item limit the conclusion? Can a narrower question be answered reliably?
Communication	Which conclusions are observations, estimates or assumptions? Can the result be explained without treating technical confidence as a legal verdict?

Together, the preservation record and expert schedule should permit disagreement to become specific. The reviewer should be able to identify whether the dispute concerns origin, alteration, method, legal relevance, causal contribution, or the consequence of missing evidence. That precision is more useful than a general declaration that an AI system is either reliable or unreliable.

4.4 Individual Review and Corrective Remedies

Starting with the injury rather than the technology

The first remedial question is what must change for the person to obtain meaningful relief. Correcting a name, removing a flag, reopening an application, restoring a payment, compensating an expense, and addressing professional misconduct are distinct outcomes. An institution may complete one while leaving the others unresolved. The chapter therefore proposes an injury-based intake process: record the adverse act, its continuing effect, the relief sought, the earliest relevant deadline, and the body capable of providing that relief.

This approach does not require the initial complaint to diagnose the whole technical chain. A claimant may be able to show that payment stopped after an incorrect match without knowing whether the cause was a model, a configuration rule, or a stale database. The complaint should describe the known event and identify what is disputed. The receiving institution can then preserve records and examine its own action. A requirement that the claimant first identify the supplier's internal defect would confuse an evidential question with the practical need for reconsideration.

Internal correction and reconsideration

An internal process can be useful when it gives a responsible person access to the relevant material and authority to change the action. Merely submitting the same inputs to the same unchanged system is not necessarily reconsideration. The review should examine the disputed fact, criterion, inference, or procedural step and determine whether the original outcome remains justified. Where the reviewer cannot depart from the output, the institution should describe that limitation rather than represent the process as independent judgment.

Correction and reconsideration should be coordinated but separately documented. The custodian may amend a source field; the deciding institution must still determine the consequence for the person's application. A correction should not disappear at the point of transfer. Brazil's LGPD provides a specific propagation rule for specified correction-related actions: article 18(6) requires communication to processing agents with whom information was shared, subject to its impossibility and disproportionate-effort qualification (LGPD, 2018, art. 18(6)). The additional question of restoring an entitlement remains tied to the applicable decision-making framework.

In the running hypothetical, an identity office confirms that two records belong to different people. The benefit administrator should then identify the decisions that relied on the erroneous link, consider the correct eligibility evidence, and determine what action it can lawfully take concerning interruption and arrears. The source custodian's email does not itself decide those matters. The proposed review record should show

the correction received, the official assessment, the resulting decision, and any further route available for disputed loss.

Internal independence should be assessed in practical terms. A reviewer may work within the same institution yet possess the necessary authority, information, and freedom to disagree. Conversely, a differently named complaints unit may have no access to the decision record or power to act. The relevant inquiry is not the unit's title but whether it can examine the issue, obtain evidence, give a reasoned response, and implement or secure an effective corrective step.

Selecting the external route

External bodies have different functions. Data-protection regulators may deal with illegal processing within their statutory remit. Sector tribunals may consider the entitlement to something or the regulation of a transaction. Professional bodies may look at the actions taken. Ombud institutions may have various investigative, advisory and/or other powers according to statute. Courts may determine if an action has been legally carried out, and decide a civil claim.

These routes will assist in identifying areas to explore - they do not mean that there is an external route for every type of AI-related dispute in each jurisdiction.

In addition to determining which category the claim falls into, the adviser should assess:

- the identity of the respondent
- the nature of the issue
- whether the claimant has standing
- what orders the competent court, tribunal or other reviewing body could make
- conditions of access to those orders.

An organization may be independent but still lack sufficient power to award the required relief. Finding that information was handled improperly does not automatically reinstate employment status, cancel a tax bill or replace lost monies due to consequential losses. Equally, simply proving that a civil claim exists does not eliminate the possibility of accessing an administrative process that could provide quicker resolution to the current deprivation.

South Africa’s Information Regulator publishes complaint procedures and identifies processing outside its remit, including processing relating to a court’s judicial functions. POPIA section 75 separately requires reasonable assistance where necessary to put a complaint in writing (Information Regulator South Africa, n.d.; POPIA, ss. 6, 75). While this documentation provides valuable information about how to file a complaint regarding improper handling of processed information, it should not be interpreted as confirmation that the Regulator can overturn every decision pertaining to use of personal information. The complainant or adviser should determine the specific

act being challenged and the specific authority providing the relief prior to making a filing.

Judicial review, exhaustion, and timing

South Africa's Promotion of Administrative Justice Act (PAJA) illustrates the importance of procedural timing. Section 7 requires review without unreasonable delay and ordinarily within 180 days. The starting point depends on whether statutory internal proceedings have concluded or no such remedy exists, in which case the specified notice or knowledge of the act and its reasons matters. Exemption from exhaustion requires an application, exceptional circumstances and the interests of justice. Sections 8-9 govern relief and extensions of time. Judicial functions excluded by section 1 do not become administrative action merely because a court uses software (PAJA, 2000, ss. 1, 7-9).

Colombia's tutela procedure illustrates a different approach. Article 6 of Decree 2591 makes the effectiveness of other judicial remedies in the applicant's circumstances relevant to subsidiarity, while preserving temporary protection against irreparable harm. Article 7 authorises provisional protective measures. Article 9 states that a prior administrative appeal is not required to seek tutela, but tutela does not dispense with exhaustion required for separate contentious-administrative proceedings (Decree 2591, 1991, arts. 6-9). These provisions show why a universal instruction to exhaust all internal remedies is inaccurate; the requirements of each route must be checked separately.

Each item in a chronological listing of a case should include the date of the final decision concerning the issue(s), notice of the adverse decision, the date when the individual becomes aware of or discovers the inappropriate use of his/her information, the date when a request is filed for records relating to the complained-of event, the date when an internal complaint is filed, and responses received, and the dates of any applications for judicial review and resulting court orders. The applicable law may attach different procedural consequences to these dates. Therefore, the adviser should verify what begins each applicable timeframe and what statutes or court decisions extend or suspend them. Delaying filing a grievance until receiving an explanation from the company responsible for collecting or utilizing your personal information is unlikely to afford you adequate protection absent that determination.

Additionally, costs associated with initiating litigation should also be identified prior to determining an appropriate course of action. Costs associated with filing charges, attorney representation, translating documents from foreign languages, traveling long distances for depositions or trials, retaining experts who testify at trial, obtaining copies of requested documents, and taking leave from your job (paid or unpaid) can make certain procedures impractical. The adviser should obtain accurate fee schedules for the applicable procedure(s) and any financial assistance available to litigants from the competent authority instead of assuming that using an online platform eliminates

costs associated with litigation. Although there may be a viable legal avenue for redress - it is equally important to demonstrate that each such alternative is practically accessible.

Limiting urgent harm

An interim response should address the harm that may occur before final resolution. In the hypothetical suspension, the immediate question may be whether continued interruption threatens an essential need, whether alternative verification is available, and which body can authorise temporary continuity. Preservation of records is necessary for investigation but does not meet that need by itself. A proposed request should identify the impending consequence, supporting facts, measure sought, duration, and legal basis.

The legal route must be specified. PAJA expressly includes temporary relief among the orders contemplated by section 8. Colombia's tutela framework provides provisional protection under article 7 (PAJA, 2000, s. 8; Decree 2591, 1991, art. 7). Neither illustration establishes that every complaint automatically suspends a decision. The applicant must address the relevant conditions, and the competent body must determine the measure and any safeguards.

The institution should consider whether a narrower measure can control the immediate risk without prejudging the entire dispute. Possible proposals include pausing further adverse reliance on a disputed record, allowing an alternative verification channel, preserving a place in a process, or maintaining a defined service temporarily where legally authorised. These are remedial design options, not universal entitlements. Their suitability depends on the injury, the legal framework, the interests of others, and the practical capacity to implement them.

Urgency should not be equated with a full technical finding. A request may be supported by a credible mismatch, a material gap in reasons, and evidence that delay will make relief ineffective. The applicant should explain why the temporary measure is needed while the disputed technical issue remains unresolved. The respondent must have the participation and protections required by the applicable procedure. Interim action should neither assume final liability nor allow unresolved supplier responsibility to become an automatic reason for inaction.

Distinguishing the available corrective outcomes

Data amendment addresses the record. Decisional reconsideration addresses the application of law to the evidence. Restoration addresses the position lost, such as reinstating access or providing a payment found to be due. Compensation addresses legally recoverable loss under an applicable cause of action. Discipline addresses professional or official misconduct. Structural correction addresses a recurring process. The proposed case record should identify which of these outcomes has been requested, which is within each body's competence, and which remains unresolved.

The distinctions also help define completion. An institution should not close the case as “corrected” merely because a field has changed. It should verify whether the original decision has been reconsidered, whether the correction reached relevant recipients, and whether the ordered or agreed relief has occurred. Where financial loss is disputed in separate proceedings, the closure record should make that boundary explicit rather than imply that every consequence has been settled.

In Colombia's tutela procedure, article 27 provides for continued judicial competence until the right is restored or the causes of the threat are eliminated (Decree 2591, 1991, art. 27). That jurisdiction-specific feature illustrates why implementation can be a distinct phase rather than an administrative afterthought. In the chapter's proposed pathway, an implementation record links the operative order or decision to the responsible institution, required action, completion evidence, and unresolved barriers.

Multiple routes can create duplication, inconsistent outcomes, and excessive expense. The response is coordination based on jurisdiction and relief, not an assumption that one route answers everything. A claimant may prioritise restoration while preserving a separate claim for loss. Institutions should avoid making a person repeat the same factual account unnecessarily, but information sharing must have a lawful basis and appropriate limits. Any agreement resolving the matter should identify what it settles and what it does not.

Choosing relief in Case M

In the South African variant, the claimant may prioritise prompt restoration while preserving a separate damages claim. The initial materials would include the suspension notice, evidence that the administrator received the identity objection, supporting identification, a request for reasons and an account of urgency. The review would examine the administrator’s reliance on the link and treatment of conflicting evidence. A High Court with jurisdiction under PAJA’s definition of 'court' is a possible forum; the wider statutory definition does not mean that every magistrate’s court automatically has jurisdiction (PAJA, s. 1).

When a statutory internal remedy exists, the adviser will identify it and utilize it or pursue the section 7(2)(c) exemption based upon meeting the exceptional-circumstances and interest-of-justice requirements. When no internal remedy exists, making

an informal request to a supplier will not transform it into an internal remedy. The 180 day parameter established in section 7(1), serves as an outer limit along with an additional requirement that there be no unreasonable delay. The triggering event varies depending upon the existence and outcome of an internal remedy and the identified knowledge of the act and reasons. Section 9 establishes permissible avenues to extend the relevant time period. Therefore, the example request cannot assure that an unanswered records request halts time (PAJA, ss. 7-9).

The requested interim measure would be articulated clearly: restore access temporarily or prevent continued reliance upon the disputed link during the review process, where legally justified while the competent body examines identity and entitlement. Any evidence offered to support the interim measure should demonstrate the impact of disrupting service; reasonable alternatives available to verify entitlement; the impact on other parties; and the scope of the order which is being requested. Authority to provide temporary relief is found in section 8(1)(e); however, utilizing this authority will require compliance with both applicable application and interdict requirements. A separate reason will also need to exist for requesting final substitution so as to establish entitlement to an extraordinary remedy. The court should not be expected to determine entitlement simply due to a demonstrated error in the data used for linking.

Implementation must be assessed against the resulting order. Correcting a link, reopening an application and paying amounts due are separate events. Under POPIA section 24(3), where a correction changes information affecting decisions about the data subject, the responsible party must, if reasonably practicable, inform the persons or bodies to whom it disclosed that information. Section 24(4) separately requires notification to the requesting data subject. The closure record should therefore establish whether relevant recipients were informed and whether the benefit decision was lawfully reconsidered and the required relief implemented. Any damages claim retains its own elements and procedural requirements (POPIA, ss. 24(3)-(4), 99).

In the Brazilian private-service variant, the claimant should distinguish correction of personal data from restoration of service. LGPD article 18 provides the correction right; article 18(6) requires immediate communication of correction, deletion, anonymisation or blocking to processing agents with whom data were shared so that they repeat the action, unless communication is demonstrably impossible or disproportionate. An application for urgent relief under CPC article 300 must address the probability of the right, danger of harm or risk to the effectiveness of proceedings, and applicable safeguards, including reversibility for anticipatory relief. Compensation under Consumer Defence Code article 14 remains dependent on a service defect, damage and causation, subject to the applicable proof rules. These routes connect correction with consequences but differ in legal basis, respondents and remedies (LGPD, art. 18; Code of Civil Procedure, art. 300; Consumer Defence Code, art. 14).

Individual remedy pathway

Table 4.4 is a planning tool for a legal adviser, legal-aid provider, complaints officer, or authorised representative. It should be completed with the actual body's name, legal basis, deadline, required form, cost, and available assistance. Entries for a regulator, tribunal, or court are conditional on jurisdiction; the table does not create a new appeal route.

Table 4.4. From complaint to effective individual relief

Stage and competent actor to identify	Essential records and procedural checks	Possible result within lawful powers	Barrier requiring attention
Initial intake: service body or adviser	Adverse act, current injury, notice, relevant dates, requested relief, authority to represent	Case acknowledgment, records preservation, routing and urgent-risk assessment	Digital-only access; no written decision; difficulty identifying the institution
Source correction: authorised custodian	Disputed field, original source, correction evidence, downstream recipients	Amendment or disputed-status entry; notification of relevant recipients	Custodian can change data but not the consequential decision
Internal reconsideration: competent decision authority	Original and corrected record, criteria, output, human action, new representations	New decision, reasons, lawful restoration or referral	Review repeats the same automated process or lacks power to depart
Independent review: regulator, sector body, tribunal or ombud where competent	Jurisdiction, standing, prior steps, complaint, requested order, disclosure needs	Investigation, recommendation or binding relief according to mandate	Limited remedial powers; exhaustion conditions; unclear transfer arrangements
Judicial proceedings: competent court	Cause of action or review ground, respondent, dates, record, standing and procedural requirements	Available interim or final order, reconsideration, restoration or damages where justified	Time limits, costs, expert access, immunities, evidential or jurisdictional disputes
Implementation and follow-up: named remedial owner and enforcing body	Operative order, timetable, recipients, completion evidence, continuing loss	Implemented correction and relief; enforcement or further review if necessary	Changed suppliers, missing records, stale replicas, unfulfilled payment or reinstatement

The pathway should be maintained as a single chronology with separate legal tracks. A named coordinator can help connect the claimant to the relevant bodies without claiming powers those bodies do not possess. Completion should be tested against the actual injury and operative decision. An acknowledgment, a corrected database, and an implemented remedy are three different achievements.

4.5 Collective Harm and Institutional Enforcement

When individual correction is insufficient

A successful individual complaint can leave the source of repeated harm unchanged. The institution may reopen one application while continuing to apply the same defective rule to others. A supplier may repair one record without identifying the affected cohort. A court may receive a corrected filing while the professional organisation continues using a workflow that produces unsupported citations. Collective enforcement begins by asking whether the evidence identifies a recurring practice that requires investigation or correction beyond the immediate case.

The relevant pattern should be defined by conduct and consequence, not merely by a shared technology. People affected by different uses of a general-purpose model do not necessarily share a legal claim. Conversely, people affected by a common exclusion rule may share a material issue even when some decisions used AI and others used conventional automation. The proposed unit of collective investigation is the contested institutional process: its criteria, data, versions, users, population, and route from output to adverse act.

Standing and representative procedures

Standing should be identified before an organisation presents itself as authorised to litigate for a population. The relevant questions include the claimant's capacity, the right or interest invoked, the relationship with affected people, and any procedural requirements for representative action. Organizational expertise or public concern does not by itself establish legal standing. Equally, a legal framework may permit a claim without requiring every affected person to appear individually.

Kenya's Constitution supplies an express illustration. Article 22 permits rights proceedings by a person acting for someone unable to act, for a group or class, in the public interest, or as an association acting for members, as well as in one's own interest. Article 23 identifies forms of relief within the court's jurisdiction, including declarations, injunctions, conservatory orders, compensation, and judicial review (Constitution of Kenya, 2010, arts. 22-23). The entitlement to bring proceedings does not dispense with proof of the alleged rights violation or establish the appropriate scope of an order.

Brazil's Public Civil Action Law identifies eligible claimants, including the Ministério Público, Defensoria Pública, specified public bodies, and qualifying associations. Association standing is subject to statutory conditions, with a defined possibility of dispensing with the prior-constitution period. The Law's subject matter and exclusions also require examination (Public Civil Action Law, 1985, arts. 1, 5). The proper comparison is therefore not that one jurisdiction "allows class actions" and another does not. It is which actor can invoke which procedure for the particular common interest and requested relief.

Representation also has a practical dimension. An organisation may possess standing while knowing little about the circumstances of some affected people. The proposed enforcement plan should identify how people are informed, how accounts are gathered, how confidentiality is protected, and how disagreement over the requested outcome is recorded. A group should not be treated as having a single preference merely because members encountered a similar technical problem. Some may seek restoration, others a non-digital route, and others protection against further disclosure of sensitive information.

Regulatory investigations and ombud institutions

Regulatory investigation can address patterns without making each affected person pursue a separate damages claim. Its availability and reach depend on the regulator's mandate, investigatory powers, procedural protections, and enforcement instruments. A request should identify the conduct within that mandate and the records capable of establishing it. The existence of a regulator does not create general jurisdiction over all consequences of an AI system.

The South African Information Regulator describes section 89 assessments that may be initiated on its own initiative or requested in the circumstances it sets out. Its published process identifies recurring complaints and security compromises as possible signals and describes documentary analysis, further evidence, and an opportunity to comment on an interim assessment (Information Regulator South Africa, n.d.). This is evidence of a published investigatory route, not evidence that any particular AI deployment has been investigated or that a recurring complaint necessarily proves non-compliance.

An ombud institution should be assessed with the same precision. The chapter's proposed inquiry asks whether it can receive a representative complaint, initiate an investigation, compel information, recommend change, issue a binding measure, or refer a matter elsewhere. These powers should be entered from the institution's actual legal instrument. A recommendation may be important even when it is not directly enforceable, but its status and implementation mechanism must be clear to the affected group.

Investigation and litigation can complement one another without being interchangeable. A regulator may obtain technical material relevant to a recurring processing practice; a court may be asked for relief outside that regulator's powers. A representative organisation may collect accounts and refer evidence to a public body without itself becoming the claimant in later proceedings. The enforcement plan should distinguish these roles and identify what can lawfully be shared between them.

Building evidence of a recurring mechanism

Aggregate evidence should begin with a defined population and outcome. What decisions are being counted? Over what period? Under which version and configuration? Does the dataset include only completed applications, or also rejected submissions and attempts that never entered the register? A count of recorded complaints cannot establish the experiences of people who could not access the complaint channel. The investigator should explain the boundary of the dataset and how that boundary affects any conclusion.

A higher adverse-outcome rate for one group is a reason to examine the process, not an automatic legal finding of discrimination. The analysis should identify the applicable legal standard, the relevant comparison, the rule or practice alleged to produce the difference, and the evidence bearing on explanation or justification. It should separate an observed association from a claim about causation and from the ultimate legal characterization. The same discipline applies when the difference is small or absent: aggregate similarity does not resolve every individual procedural or rights claim.

The running hypothetical can be extended to illustrate a bounded investigation. Several applications have been suspended following duplicate-record flags. The investigator would examine whether the same matching rule was used, whether errors cluster around a particular source or date, and whether officials had an effective exception route. A finding that names were transliterated differently would lead to one inquiry; a rule treating every unresolved match as ineligibility would lead to another. The remedy should address the demonstrated mechanism rather than assume that retraining the model will resolve both.

The investigation should consider alternative explanations openly. A change in the population applying, eligibility requirements, staffing, source-record quality, or complaint publicity may affect observed outcomes. A before-and-after comparison that ignores these changes can attribute too much to a model update. Where the evidence does not support a causal estimate, the report can still document a common process, recurring errors, and the consequences of unavailable records without claiming more than has been established.

Quantitative evidence should be connected with case reconstruction. Individual files can reveal how a category became an adverse act, whether contrary evidence was available, and what happened when a person sought correction. Aggregate data can indic-

ate how broadly the identified mechanism may operate. Neither should automatically replace the other. A striking case does not establish prevalence, while an average can conceal an important exception or an inaccessible pathway affecting a smaller group.

Expert access, confidentiality, and affected-group participation

A system-directed claim may require access beyond the records of one person. The request should identify the relevant cohort, variables, time period, configuration history, and sampling method. It should explain how the requested information bears on a common issue. A request for all institutional data without a defined purpose can create unnecessary privacy and security risks while producing material too broad to examine effectively.

Brazil's Public Civil Action Law provides for civil inquiries by the Ministério Público and requests for information and expert material, with specific treatment of legally confidential information. It also contemplates obligations to act or refrain from acting and provisional orders (Public Civil Action Law, 1985, arts. 3, 8, 11-12). These provisions illustrate a legal vehicle for investigation and relief within the Law's scope; they are not an unrestricted entitlement for any organisation to inspect any private model.

Proposed access arrangements should protect both effective scrutiny and the people represented in the records. An independent expert may need identifiable linkage to reconstruct a false match, while a public report may require only carefully limited findings. The plan should distinguish investigative access from publication. It should also record who approves the access, who can challenge its scope, and what happens to the material when the inquiry ends.

Affected-group participation should extend to the design and evaluation of the remedy. An institution may propose an online reapplication exercise that appears administratively efficient but does not reach people whose difficulty was inability to use the original channel. The relevant question is whether the people covered by the corrective measure can actually benefit from it. Consultation should be documented as a process with limits and disagreements, not used as a substitute for the legal basis or evidence supporting the order.

Temporary safeguards and structural correction

A temporary safeguard should be proportionate to the documented risk and authorised through the relevant procedure. It might involve additional review of a disputed category, an alternative intake route, preservation of an existing service, or suspension of a defined automated step. The proposal should explain why that measure addresses the possible harm and how essential functions will continue. Suspending an entire service without a continuity plan can create a different problem for the same population.

Structural correction should identify the change required, the responsible actor, and the affected decisions. Possible measures for consideration include revising a configur-

ation, correcting source records, changing an exception procedure, reprocessing an identified cohort, improving notices, or providing necessary review access. Each measure must be linked to the established legal defect and the decision-maker's lawful powers. A generalized direction to “remove bias” or “make the algorithm transparent” is difficult to implement or assess without specifying what must change.

T-323/24 illustrates a distinct institutional response: the Colombian Court's orders concerning guidance and training went beyond its resolution of the particular judicial-AI issue. Those orders do not establish that compliance later occurred (Corte Constitucional de Colombia, 2024, operative orders five and six). For the present framework, the important question is how a general instruction becomes an implemented practice whose operation can be examined.

The institution should not assume that a new model version repairs earlier decisions. Prospective correction and retrospective relief require separate plans. The first addresses future use; the second identifies affected people, determines whether their decisions require reconsideration, and implements the relief available. Reprocessing should use the legally relevant criteria and evidence for the matter, not simply substitute a current model for a historical one and accept whichever answer it produces.

Monitoring compliance and defining completion

A structural measure needs an observable implementation record. The plan should specify what evidence will show that the required change was made and that the corrective route works. Relevant indicators may include completion of identified reviews, correction reaching recipients, delivery of notices, availability of an alternative channel, and implementation of payments or reinstatements found to be due. These are proposed monitoring subjects, not invented baseline values or universal targets.

Monitoring should distinguish outputs from outcomes. A revised policy, training session, installed update, or closed ticket is an output. The ability of an affected person to obtain reconsideration, correction, or restored access is a different outcome. A report should explain the relationship between the two. It should also identify unresolved cases, unsuccessful attempts, exceptions, and disagreement rather than treating administrative closure as proof of substantive compliance.

System-wide measures can overreach if their scope exceeds the established wrong or the institution's ability to implement them lawfully. This concern supports a defined population, staged implementation, opportunities for the respondent to address technical and operational constraints, and review of the measure's effects. It does not justify limiting every response to the person who first complained. The appropriate boundary is the relationship between the proven or properly established concern, the legal power invoked, and the corrective measure.

An exit from enhanced supervision should also be reasoned. The responsible body should identify the evidence supporting completion, the matters still reserved, and the

conditions that would trigger renewed scrutiny. Where a supplier is replaced during the process, the original institution should retain access to the records needed to implement existing obligations. Changing a contractor should not be treated analytically as proof that the underlying practice has been corrected.

Collective enforcement plan

Table 4.5 is intended for a competent regulator, litigation team, public-interest organisation, or institutional reviewer. It does not presume that these actors have identical powers. The completed plan should identify who investigates, who represents affected people, who can issue an order, who implements it, and who verifies compliance.

Table 4.5. Investigation, structural correction, and monitoring

Phase	Evidence and participation required	Measure requiring a competent legal basis	Completion or escalation test
Define scope and authority	Common practice, population, period, claimed right or duty, standing and institutional mandate	Open the appropriate investigation or proceeding	Claim and requested relief fall within the chosen route
Preserve and investigate	Case sample, aggregate records, versions, criteria, complaints, source history, expert access	Proportionate preservation and production; protected access where available	A recurring mechanism is established, excluded, or remains specifically unresolved
Protect against continuing harm	Urgency, affected interests, alternative channels, respondent's explanation and service dependencies	Temporary review, continuity, restriction or other authorised safeguard	Risk is limited without creating an unexplained interruption of essential service
Design correction	Findings, source and configuration defects, legal criteria, affected-group input	Defined changes to rules, records, procedure, notices, review or deployment	Each required change has an owner, timetable and testable result
Address past decisions	Identifiable cohort, historical versions, correction evidence and individual circumstances	Reconsideration, notification, restoration or other relief within lawful powers	People covered by the measure receive the review and relief determined to be due

Table 4.5. Investigation, structural correction, and monitoring

Phase	Evidence and participation required	Measure requiring a competent legal basis	Completion or escalation test
Monitor and close	Independent checks, unresolved cases, completion records, costs, access and continuing complaints	Reporting, further directions, enforcement or termination of supervision as authorised	Compliance is demonstrated rather than inferred from a policy or software update

The plan should preserve the distinction between an investigation signal and a legal finding. Recurring complaints can justify scrutiny; a confirmed common error can justify a more focused corrective response; a system-wide order requires the relevant legal foundation and evidence for its scope. Those stages should remain visible in reports and public communications.

Integrating the challenge-and-remedy pack

The chapter's practical outputs should operate as connected records. The responsibility matrix identifies actors and duties; the notice and request templates identify the act and disputed information; the preservation checklist and expert schedule support reconstruction; and the individual and collective pathways connect findings to relief. A shared reference should link these records without requiring every participant to receive every confidential item.

For an individual matter, the pack should make it possible to answer who decided, what was relied upon, what is disputed, which evidence is preserved, which body can act, and whether the resulting relief was implemented. For a recurring practice, it should add the population and version boundaries, representative authority, common mechanism, and structural obligations. This connected documentation is the practical condition for preventing a technically complex dispute from becoming a sequence of referrals in which no actor addresses the injury.

Chapter 4 Key Arguments

Responsibility begins with a specified act and an identified legal duty. The paired false-match example shows that information correction, administrative validity and compensation can require different respondents and proof, even where they arise from one event.

Disclosure and evidence become useful when tied to a disputed inference. Preserving a file can establish what an institution received or communicated without establishing the truth of its contents. An error can justify reconsideration without establishing a damages claim against every participant.

The remedy record must follow the claimant from the initial adverse act to implemented relief. Jurisdiction, procedural time, access costs and the authority to implement an order shape that pathway as much as the technical explanation.

The practical tools are therefore records of legal reasoning and coordination. They cannot supply missing jurisdiction, create a cause of action or establish effectiveness simply by being completed.

CHAPTER

05

Regulation, Procurement, and Institutional Implementation

IN THIS CHAPTER

5.1 Regulatory Instruments and Institutional Mandates

5.2 Procurement Requirements and Contractual Safeguards

5.3 Impact Assessment and Local Validation

5.4 Staffing, Training, and Financial Planning

5.5 Evaluation, Incident Response, and System Withdrawal

Introduction

An institution does not acquire the capacity to govern artificial intelligence simply by purchasing a system, publishing a policy, or appointing a committee. It acquires that capacity only when particular people can establish the permitted use, examine the relevant evidence, control changes, respond to affected persons, and maintain a lawful service when the technology fails or is withdrawn. The practical question is therefore not whether a governance framework exists, but whether its requirements can be carried out under the institution's actual conditions of work.

The preceding chapters approached legal transformation through authority, professional practice, information, decisions, and remedies. This final chapter brings those inquiries into the arrangements through which a system is selected and operated. A legal limit identified during initial assessment must become a restriction in the specification and operating procedure. A data-quality concern must become a test, a correction responsibility, and a retained record. A right to challenge a decision must be supported by access to evidence and an official able to reconsider the outcome. The possibility of withdrawal must be translated into contractual rights, usable exports, trained staff, and arrangements for unresolved cases. Otherwise the book's earlier safeguards remain disconnected statements rather than an institutional method.

The African Union's Continental Artificial Intelligence Strategy provides one starting point. Its governance discussion calls for examination of existing legal frameworks and regulatory gaps, while its implementation section distinguishes national responsibilities from regional coordination and support. It recognises variation in members' circumstances and resources; its implementation proposals are not evidence that particular institutions have acquired the necessary powers or capabilities (African Union Commission, 2024, §2.4.1, pp. 32-33; §§3.1-3.2, p. 60). UNESCO's official description of its 2025 Guidelines for the Use of AI Systems in Courts and Tribunals likewise emphasises human oversight, rights, accountability, and an assistive rather than substitutive role for judicial AI. That description supplies an orientation, not a domestic rule authorising any particular use (UNESCO, 2025, publication description).

The chapter develops an implementation handbook for the book's focal settings: Pakistan, India, Kenya, South Africa, Brazil, and Colombia. Its comparisons remain selective. A procurement rule, a data-protection statute, a judicial resolution, and a constitutional judgment are examined for the particular institutional work each can perform. They are not interchangeable sources of authority, and an arrangement documented in one jurisdiction is not presumed to exist elsewhere. The absence of an example for a particular country or function signifies a limit to the selected evidence, not the absence of law or institutional activity.

The main argument is that implementation depends on connected, reviewable decisions rather than a single approval. Those decisions concern the choice of legal instrument, the terms of acquisition, the evidence supporting a pilot, the allocation of

continuing work, and the response to subsequent findings. The proposed tools make each decision attributable to an actor and connect it to records that another person can examine. They also leave room for a conclusion that the proposed use is not justified, that a simpler alternative addresses the problem, or that the institution cannot yet operate the necessary safeguards.

5.1 Regulatory Instruments and Institutional Mandates

Matching the instrument to the institutional problem

The expression “AI regulation” can conceal several different questions. One concerns who may perform a legal function. Another concerns the conditions attached to processing information. A third concerns the conduct of a profession, the management of a court, or the purchase of a service. Yet another concerns who can investigate a failure and require a response. A single project can engage all of these questions without placing them within the competence of a single body. Institutional design begins by separating the question to be answered from the instrument being considered.

Legislation can establish duties, confer powers, define institutional relationships, and provide enforcement routes within the relevant constitutional order. Sector regulation can specify obligations within an existing enabling framework. Court rules and judicial governance instruments address matters within the issuing institution’s competence. Professional standards govern the conduct of covered practitioners through the applicable professional arrangements. Procurement documents and contracts specify what a purchaser requires from a supplier. Guidance can explain an approach, support consistency, and assist interpretation, but its legal effect depends on its source and any valid incorporation. None should be assigned the functions of another merely because its wording is detailed or mandatory in tone.

Consider a hypothetical court registry proposing an intake classifier and a translation assistant. The classifier would suggest a queue; the translation assistant would help staff understand submissions. The institution needs to establish its authority to organize intake, the limits on changing filing requirements, the conditions for sending documents to a provider, the status of translated material, and the responsibility for urgent matters. A vendor contract can require an exportable queue history. It cannot itself establish the registry’s power to reject a filing or change a statutory deadline. Conversely, a rule requiring review is not operationally complete if the institution cannot retrieve the file that the reviewer must examine.

Existing bodies and the boundaries of their powers

An institution's suitability for a particular oversight task has at least four dimensions: legal competence, relevant expertise, independence in performing that task, and resources to act on the resulting findings. These dimensions should be recorded separately. A body may have power to obtain documents but insufficient specialist support to interpret them. Another may understand the technology yet lack authority to compel cooperation. A third may investigate complaints but not reverse the underlying decision. The assessment must not turn one demonstrated capacity into an assumption about the others.

South Africa's Protection of Personal Information Act illustrates a defined statutory allocation. Section 40 gives the Information Regulator functions involving education, monitoring and enforcement, assessment, and complaints concerning personal-information processing. Those functions arise within the Act's scope and conditions. They do not make the Regulator a general appeal body for every substantive outcome in which AI appears (Protection of Personal Information Act 4 of 2013 [POPIA], s. 40). The analytical distinction is between oversight of processing and reconsideration of the service decision, even where the same incident requires both.

Kenya's Data Protection Act similarly allocates oversight and investigation functions to the Office of the Data Protection Commissioner, provides information-gathering powers, and permits delegation to a statutory regulator subject to conditions imposed by the Data Commissioner under section 10. This is an example of coordination being given a legal form, not proof that every overlapping sector has an operating joint arrangement (Data Protection Act 24 of 2019, ss. 8-10). A proposed partnership must still identify what is delegated, which body receives the evidence, and who remains responsible for the statutory task.

Other oversight questions require different institutional routes. Financial scrutiny asks whether expenditure was authorised and supported, rather than deciding a claimant's eligibility. Professional supervision asks whether an adviser performed the work required of that role. A court considers the matters placed within its jurisdiction through applicable procedures. Internal management can direct staff and control authorised operations, but should not be confused with independent review of its own conduct. The chapter's proposed map therefore records the action each body can take, not merely its participation in a committee.

This distinction matters to affected persons. An individual should not have to infer whether the error is a privacy breach, a classification defect, a professional failure, or an unlawful refusal before receiving assistance. An intake office can accept and route the complaint while preserving the different legal processes. Such coordination is an administrative proposal: it does not extend statutory standing, limitation periods, or powers of relief. The institution must explain which matters have been transferred, which remain pending, and whether a separate filing is needed to protect a deadline.

Judicial governance as a distinct institutional setting

Courts present a particular combination of adjudicative independence, administrative management, and professional responsibility. A centralised technical service may support several courts without acquiring authority over the merits of cases. An administrative approval to use software should therefore identify whether it concerns research support, registry work, translation, or a contribution to reasoning. The governance instrument must be read alongside the allocation of judicial functions and applicable procedural protections rather than treated as a complete code for all judicial conduct.

The paired judicial comparison in section 2.3 traces Colombia's T-323/24 through the Superior Council's Acuerdo PCSJA24-12243 of 16 December 2024, and examines Brazil's Resolution 615/2025 as a different governance instrument. The important implementation distinction is between an order to make rules, the adoption of those rules, and evidence that institutions comply with them. Adoption supplies a legal fact; operational compliance requires a further evidential account (Corte Constitucional de Colombia, 2024, operative order five; Consejo Superior de la Judicatura, 2024; CNJ, 2025).

The comparison also limits the role of the contract. A judicial instrument can establish the institution's obligations within its lawful field; procurement can secure records, testing access and assistance needed to perform them. Neither a supplier undertaking nor an operating manual can confer judicial competence. The approval record should connect each requirement to its source and name the institutional actor who must act when contractual delivery proves insufficient.

What a new legal instrument would need to change

The inquiry about legislation should identify an actual legal gap. Possible gaps include an absence of power to require a supplier's records, uncertainty over an institution's competence, a missing duty to preserve particular evidence, or the lack of a procedure through which affected persons can seek relief. Each alleged gap requires an account of the existing law and why its available mechanisms do not address the specified problem. Technical novelty, by itself, does not establish such a gap.

Other deficiencies are operational. Unfilled positions, incomplete records, weak translation support, or a failure to implement an existing review procedure do not necessarily indicate missing legislative authority. They may concern funding, management, contractual terms, training, or compliance with a duty already in force. The distinction matters because changing the legal text does not demonstrate that the institutional work can be performed. Equally, a well-resourced project team cannot create a coercive power that the applicable law does not supply.

The comparison between a new institution and an existing one should therefore remain functional rather than categorical. A proposal for a new body would need to specify its jurisdiction, relationship to sector regulators and courts, information powers,

reviewability, financing, and approach to conflicts. A proposal relying on existing bodies would need an equally precise account of allocation, referrals, shared expertise, and unresolved responsibilities. Neither institutional form establishes effectiveness in advance, and this chapter does not rank them or infer their performance from organisational design alone.

Coordination, regional cooperation, and foreign models

Coordination becomes concrete when bodies agree, within their powers, on which records to request, how to preserve them, how to avoid inconsistent instructions, and how to inform an affected person of the applicable route. A coordinating office can maintain a system inventory or arrange technical support. That does not necessarily entitle it to approve substantive decisions, decide complaints outside its jurisdiction, or disclose protected information. Shared work requires a lawful information arrangement as well as an administrative agreement.

Regional cooperation can support shared terminology, training, testing methods, and exchanges about observed failures. The transferable object might be an evidence-request format or a method for testing a translated deadline, rather than another country's statutory threshold. Cooperation also needs attention to language, confidentiality, competing professional obligations, and the resources available to receiving institutions. A shared laboratory's report should identify the version and setting tested; it should not become a portable authorisation for materially different uses.

Foreign procurement materials illustrate the same distinction. The 2025 resources on the Commission-supported Public Buyers Community include AI model contractual clauses and explanatory commentary. The commentary expressly treats the clauses as supplements to a wider contract rather than a complete agreement, leaving important matters such as payment, acceptance, governing law, and liability to the surrounding arrangement (Public Buyers Community, 2025, commentary, p. 2). The useful comparative question is what documentary or operational problem a clause addresses. Importing its wording without checking local power, enforceability, terminology, and cost can leave that problem unresolved.

An institution adapting an external framework should prepare a short correspondence record. For each proposed requirement, it should state the foreign source, the local purpose, the local legal basis if one exists, the body expected to act, the evidence required, and the adaptation made. Where no local duty corresponds to the foreign requirement, the record should say whether the institution proposes a contractual condition, an internal standard, or further legal consideration. This prevents borrowed language from being presented as law already applicable to affected people.

Regulatory options matrix

Table 5.1 compares the work that different instruments can perform. The implementing-body entry should identify the precise enabling provision and competent office, as well as any necessary approval or consultation. The enforcement entry should distinguish the power to investigate from the power to order correction or provide relief. Operational and financial reviewers should verify the resources listed; an unresolved authority or resource assumption is a reasoned restriction on the proposed use.

Table 5.1. Regulatory instruments, implementing bodies, and enforcement questions

Legal vehicle	Implementing body and precise legal authority	Resources and evidence needed	Enforcement mechanism and limit to establish
Primary legislation	Competent legislature; institutions assigned functions	Legal drafting, institutional mapping, implementation and recurrent-resource analysis	Defined duties, information powers, remedies and review; actual scope and commencement
Sector regulation	Regulator or rule-maker acting under an enabling provision	Sector expertise, case records, technical assistance and inspection capacity	Directions, licensing measures or sanctions only within conferred powers
Court rule or judicial governance instrument	Court or judicial body with the relevant competence	Judicial and registry expertise, procedural records, training and protected technical support	Procedural or administrative controls; distinction from deciding individual merits
Professional rule or standard	Competent professional body and supervisors	Task-specific competence assessment, confidential work review and disciplinary resources	Supervision or discipline under applicable arrangements; no assumed public enforcement power
Procurement conditions and contract	Authorized purchaser and contract manager	Clear specification, independent testing, records access, change and exit capacity	Acceptance, correction, suspension, termination or contractual relief subject to law and agreed terms
Internal operating instruction	Institution and official authorised to direct the work	Named staff, versioned procedures, escalation and implementation records	Internal scope and compliance controls; no creation of external powers or removal of rights

Table 5.1. Regulatory instruments, implementing bodies, and enforcement questions

Legal vehicle	Implementing body and precise legal authority	Resources and evidence needed	Enforcement mechanism and limit to establish
Nonbinding guidance or shared technical resource	Issuing institution or cooperation arrangement	Maintained guidance, linguistic and technical support, transparent source status	Persuasive or incorporated effect to be established; no assumed coercive force

The matrix should be accompanied by a brief explanation of unresolved interfaces. It is not enough to state that privacy is assigned to one regulator and professional conduct to another. The institution must describe what happens when the same event concerns both, which evidence will be preserved, and how a person's urgent needs will be handled while responsibility is clarified. Regulatory design becomes implementable when these interfaces are supported by procedures and records, rather than left to informal cooperation after a dispute has arisen.

5.2 Procurement Requirements and Contractual Safeguards

Defining the need before specifying the service

Procurement begins with a description of the work that needs to change. A public or professional institution should identify the existing process, its documented difficulty, the population affected, and the outcome sought. "Introducing AI" describes an acquisition preference rather than a need. A more useful statement would identify the inability to locate current authorities, delays in recognising urgent submissions, or the cost of translating a defined class of notices. The comparison should include records improvement, additional staff, clearer forms, ordinary search, and simpler automation where these could address the same difficulty.

Pakistan's federal procurement example requires a temporal distinction. PPRA's official online text, checked on 20 October 2026, identifies the Public Procurement Rules, 2026 as notified by S.R.O. 1658(I)/2026 on 28 September 2026. Rule 1 provides immediate commencement; rules 68-69 preserve the Public Procurement Rules, 2004 for proceedings initiated before commencement while otherwise repealing them. Within the 2026 framework, rule 15 addresses planning, rule 17 specifications and testing, and rules 41-42 disclosed evaluation criteria and evaluation against the bidding documents. Rule 3 concerns federal procuring agencies, rather than every provincial or local purchase. These provisions support defining relevant AI requirements before competition; they do not prescribe every safeguard proposed here (Public Pro-

curement Rules, 2026, rr. 1, 3, 15, 17, 41-42, 68-69). The account uses the regulator's online presentation; the separate Gazette notification was not independently retrieved.

A specification should describe the complete service boundary. This includes the system's task, users, data sources, working languages, proposed outputs, decision points, dependencies, and excluded uses. The hypothetical registry classifier, for example, might be authorised to suggest routing while leaving all submissions visible to an officer. A product demonstration that automatically closes an incomplete file would then demonstrate a different workflow, not simply a more efficient version of the specified service. The tender should enable the institution to recognise and reject that substitution.

Before signature, the purchaser also needs to know which requirements the supplier can actually perform. A reseller may not control the underlying model's update cycle, retention settings, or technical documentation. A promise that depends on another provider's cooperation requires evidence of the relevant upstream arrangement. Where necessary access cannot be secured, the options include a narrower use, another technical arrangement, or not proceeding with that service. A contractual sentence unsupported by the supplier's rights and capabilities is not an adequate substitute for evidence.

Acceptance, evidence, and the limits of general assurances

Acceptance should concern the configuration that will be used, not a product family or a general benchmark. The record should identify the test material, local conditions, performance measures, reviewers, unresolved findings, and the version accepted. Testing should include failures and exceptions: a submission in an unsupported language, an unreadable attachment, a request for urgent relief, a missing identity field, a contradictory source, or a temporary loss of connectivity. The question is what the service does when normal assumptions fail, and whether a responsible person can intervene before an adverse consequence follows.

Security, confidentiality, and data-use obligations require separate attention. A supplier may be permitted to process a document for the institution's task without being permitted to use the document, annotations, feedback, or derived representations to improve a general service. The contract should distinguish these purposes and specify which copies and environments are covered. In South Africa, where POPIA applies, section 21 requires a written contract addressing an operator's maintenance of the security measures contemplated by section 19. That statutory illustration does not establish that a generic confidentiality clause satisfies every processing obligation (POPIA, ss. 19-21).

The following clauses are original drafting models, not a complete contract or a statement of mandatory law. They assume that the purchaser has separately established its authority to undertake the activity. "Institution" and "Supplier" must be replaced by

the correct legal parties; the schedules must contain actual technical, financial, and legal particulars. Public procurement constraints, mandatory terms, liability law, dispute procedures, and the rights of affected persons require local review. A clause should not be adopted until its promised evidence can be obtained and its operational cost has been identified.

Annotated model clauses

Clause 5.2.1. Defined service, approved configuration, and excluded uses

Model wording. The Supplier shall provide only the functions identified in the Service Schedule, using the components and configurations recorded in the Approved Configuration Register. The Schedule shall identify intended users, authorised data categories, permitted outputs, working languages, and excluded uses. No output shall initiate an external action except through the approved workflow. The Supplier shall identify any feature that cannot be disabled or separated from an excluded use before acceptance. Expansion of purpose, population, decision role, or connected service requires written approval through the agreed change procedure.

Annotation. This clause makes the subject of acceptance identifiable. A product name alone is insufficient because interfaces, retrieval collections, prompts, thresholds, or permissions can change the practical function. The schedule should distinguish assisting an officer from determining entry, priority, eligibility, or outcome. Local counsel must identify who can approve each expansion and whether it requires a different legal instrument or procurement procedure. Configuration control may increase integration and support costs; those costs should be priced rather than concealed behind an unlimited description of "customization."

Clause 5.2.2. Institutional information and prohibited secondary use

Model wording. The Supplier shall process Institutional Information solely for the documented purposes and operations in the Data Schedule. It shall not use that information, including identifiable prompts, attachments, annotations, feedback, logs, or derived representations, for general model training, unrelated evaluation, sale, advertising, or another customer's service without a separate written instruction supported by the Institution's lawful authority. The Supplier shall disclose necessary safety or security processing, recipients, retention, and technical limits before processing begins. No inconsistent online term shall expand the permitted use.

Annotation. The clause separates service delivery from development of a provider's wider product. It does not assume that the institution owns every document or can authorise any reuse it chooses. Licensing, confidentiality, personal-information rules, and the rights of third parties remain relevant. Necessary abuse detection or service security should be described narrowly rather than prohibited in a way that makes the service impossible to operate. The parties must confirm precedence among the signed

agreement, incorporated schedules, and provider terms. A dedicated environment or restricted retention arrangement may have a different price.

Clause 5.2.3. Provenance, dependencies, and accuracy of supplier disclosures

Model wording. The Supplier shall maintain a dated inventory of material models, datasets supplied with the service, retrieval collections, software components, providers, and operational dependencies. It shall identify the information available about provenance, permitted uses, known limitations, and the basis of performance claims. Where a requested fact is unknown or unavailable to the Supplier, it shall state that limitation and its consequences. Material disclosures relied upon for award or acceptance shall be contractually recorded. The Supplier shall notify the Institution when it learns that a relied-upon disclosure was materially inaccurate or incomplete.

Annotation. An obligation to disclose uncertainty is more useful than an unsupported warranty that every training source is known and lawful. The purchaser should determine which unknowns are compatible with the proposed use and which prevent adequate scrutiny. This clause does not confer ownership of a foundation model or a general right to its entire training corpus. Evidence requirements should be proportionate and connected to the service. Verifying provenance, licences, and dependencies consumes specialist time; the pricing and responsibility for that work should be explicit.

Clause 5.2.4. Local validation and conditional acceptance

Model wording. Acceptance shall require completion of the approved Test Schedule by reviewers designated by the Institution with the independence specified in that Schedule. Testing shall cover the intended legal task, relevant languages and access needs, data-quality conditions, exception handling, and the operation of human controls. The Supplier shall provide the agreed testing access and shall not represent a demonstration or an earlier evaluation as acceptance of a different configuration. Unresolved material findings shall be recorded. Live use outside the expressly approved pilot scope is prohibited until the competent institutional signatory records acceptance.

Annotation. The test schedule should identify measurable outcomes and disqualifying conditions before results are known. It should also state who resolves disagreement between legal and technical reviewers. No universal accuracy percentage follows from the clause. A search task, a translation task, and an urgency classifier need different criteria. Independent testing may require additional environments, restricted datasets, interpreters, and repeat runs. The contract should allocate these costs and specify whether retesting after a supplier-caused failure is included. Acceptance cannot legalize an otherwise unauthorised activity.

Clause 5.2.5. Decision records, retention, and preservation

Model wording. The Supplier shall create and deliver the records specified in the Evidence Schedule in documented, searchable, and exportable formats. Records shall identify the material input sources, operative configuration, output, relevant human intervention, time, and associated transaction or case identifier to the extent necessary for the approved purpose. The Schedule shall specify custody, integrity controls, access, retention, and disposal. A lawful preservation instruction shall suspend scheduled destruction for the identified records. Preservation shall not authorise unrelated reuse or unrestricted access. The Supplier shall record any material evidence that the service cannot generate or retain.

Annotation. The schedule must distinguish procurement documents, operational records, security logs, and evidence relevant to a particular dispute. It should not require indefinite storage of every sensitive interaction. Retention duties, privacy requirements, archival rules, and legal holds must be reconciled under local law. Export should be tested for meaning, not only file availability: a table of identifiers without the relevant field definitions may be unusable. Storage, secure access, indexing, and retrieval costs should be estimated for the required period, including after ordinary service delivery ends.

Clause 5.2.6. Independent examination and lawful disclosure

Model wording. The Supplier shall cooperate with examinations by the Institution and reviewers lawfully authorised for the agreed purpose. Cooperation shall include the scheduled documentation, relevant personnel, controlled testing access, and assistance interpreting retained records. The Supplier shall support production of relevant material required through applicable judicial, regulatory, audit, or other lawful procedures. Confidential material shall be handled through proportionate access controls and protective arrangements. The Supplier shall identify a specific legal or third-party restriction promptly and assist in obtaining an alternative form of access; it shall not treat a general confidentiality label as a complete response.

Annotation. An audit right needs an identified reviewer, scope, method, and delivery obligation. Access to decision records may suffice for one inquiry, while a disputed technical mechanism may require more detailed inspection. Source-code access is not automatically necessary or sufficient. The contract cannot grant an institution powers over non-parties that the law does not provide, nor override a lawful restriction. Upstream licences and support agreements should therefore be checked before signature. Controlled inspection, expert assistance, redaction, and litigation support can create substantial costs and should not be left to ad hoc negotiation after a dispute.

Clause 5.2.7. Subcontractors, locations, and continuing responsibility

Model wording. The Supplier shall identify each material subcontractor and provider, the work performed, the information accessible to it, and relevant processing or support locations. No material substitution or relocation shall occur without the notice and approval required by the Dependency Schedule. The Supplier shall secure corresponding obligations from its subcontractors for authorised use, security, evidence, incident cooperation, continuity, and exit. Subcontracting shall not reduce the Supplier's agreed performance obligations. The Supplier shall disclose any dependency for which equivalent obligations cannot be obtained before the Institution relies on that dependency.

Annotation. The clause addresses the chain of delivery without assuming that all actors are independently liable to the institution or to affected persons. Local transfer restrictions, confidentiality rules, and access procedures need separate analysis. Remote support can matter even where storage remains local. A supplier's corporate group should not be treated as one legally interchangeable actor. Detailed chain management may limit available commercial services; that limitation should inform the proposed scope and price rather than be hidden in a promise of unrestricted subcontracting.

Clause 5.2.8. Material changes and emergency maintenance

Model wording. A material change includes a change reasonably capable of affecting the permitted purpose, legal criteria, data uses, error consequences, human control, evidence access, security, or continuity. The Supplier shall provide the change description, expected effects, testing evidence, and rollback or alternative-service arrangements before the agreed approval point. Routine maintenance shall follow the separate maintenance procedure. An urgent security measure may be applied through the documented emergency process where delay would create greater exposure; the Supplier shall preserve the change record and support prompt review. Emergency maintenance does not authorise a new substantive use.

Annotation. A blanket prohibition on updates can leave a system insecure, while unrestricted updates can invalidate acceptance. The contract should distinguish change categories and specify the official authorised to act in an emergency. Where a hosted provider cannot restore a previous model version, that limitation must be disclosed. Continuity may then need to rely on another service or a manual process rather than an unsupported promise of rollback. Testing and transition costs should be allocated, including when a change originates in a provider outside the immediate supplier's control.

Clause 5.2.9. Incidents, notification, and corrective cooperation

Model wording. The Supplier shall notify the designated institutional contact of a reportable event within the period stated in the Incident Schedule and provide updates as material facts become known. Events include unauthorised information use, material security compromise, unexplained loss of evidence, serious or repeated consequential error, unapproved change, and interruption of required safeguards. The Supplier shall preserve relevant evidence, support containment and investigation, identify affected versions and transactions, and assist corrective action and required notifications. Contractual reporting shall not replace either party's applicable statutory obligations or require confirmation of final causation before an initial report.

Annotation. The reporting period must be set in relation to applicable law, service consequence, and realistic detection capacity. This chapter supplies no universal legal deadline. A technically functioning system can produce a rights-related incident, so the schedule should not be restricted to outages or intrusions. The initial report may be incomplete but should distinguish established facts from provisional concerns. The parties should specify who funds urgent assistance and how costs are later allocated without delaying protection. The reporting procedure should distinguish an initial incident report from any admission of legal liability.

Clause 5.2.10. Remedial assistance and allocation of financial exposure

Model wording. The Supplier shall provide the agreed assistance to identify and correct supplier-controlled defects, reconstruct affected outputs, support reconsideration by the competent institution, and implement lawful remedial instructions within its contractual role. The Liability Schedule shall state the warranties, exclusions, caps, indemnities, insurance, and remedies applicable to defined categories of loss. Any service-credit mechanism shall identify the failures it addresses and whether it is exclusive for those failures. No term shall purport to determine the statutory rights of a non-party or relieve the Institution of a public or professional duty that applicable law retains with it.

Annotation. The purpose is to make risk allocation visible, not to prescribe unlimited liability or a universal indemnity. The institution should examine supplier solvency, available insurance, exclusions, jurisdiction, and realistic recovery. A promise of compensation has limited practical value if no recoverable assets or coverage support it. A financial cap does not answer who can restore a benefit or correct an official record. The clause should therefore separate operational remedial work from disputes about final financial responsibility, with local advice on enforceability and any mandatory limits.

Clause 5.2.11. Portability and essential-service continuity

Model wording. The Supplier shall maintain and test the continuity and export arrangements specified in the Continuity Schedule. Exports shall include the Institution's authorised records, associated metadata, configurations and documentation needed for the agreed transition, subject to identified rights and restrictions. The Supplier shall provide assistance necessary to verify completeness and usable interpretation. Essential-service arrangements shall identify the functions maintained during interruption, the responsible parties, and the means of routing unresolved matters. The Supplier shall not make ordinary access to retained institutional evidence dependent solely on a live production subscription.

Annotation. Portability is not a promise that a replacement will behave identically. An export may preserve records without reproducing a proprietary model. The institution must distinguish what it needs to investigate earlier decisions from what it needs to continue future work. Escrow, alternative hosting, or specialist transition support may be appropriate for some services, but not universally feasible. Recovery and manual procedures require testing under realistic demand. The price should include export assistance and specify additional charges rather than allow an undefined exit bill to determine whether withdrawal is practical.

Clause 5.2.12. Suspension, termination, and controlled transition

Model wording. The Institution may restrict or suspend the agreed service in the circumstances specified in the Suspension Schedule, through its authorised signatory and subject to applicable law. The contract shall distinguish urgent protective restriction from termination for breach and any separately agreed termination right. On exit, the Supplier shall provide transition assistance, preserve records subject to lawful holds, return information in the agreed formats, revoke access when instructed, and verify disposal when legally due. Termination shall not end surviving obligations concerning evidence, confidentiality, agreed assistance, or unresolved contractual claims.

Annotation. The schedule should identify evidential triggers, notice and cure provisions where applicable, the scope of urgent action, and who decides whether use can resume. Not every defect warrants termination, and cessation of service does not resolve earlier decisions. Public procurement, insolvency, funding, and dispute rules may constrain the available arrangements. The institution should establish the authority to pay for transition and any continuing service before signing. Deletion must follow, not defeat, preservation duties; a destruction certificate should describe its coverage and exceptions rather than make an unverifiable promise about every historical copy.

The pre-signature decision record

The clauses work only when their schedules are completed. Before signature, the institution should hold a defined use record, supplier and dependency disclosures, an agreed testing method, evidence and access arrangements, a lawful data-use schedule, change controls, incident contacts, continuity arrangements, and a priced exit mechanism. Matters deliberately deferred should have an owner, a completion condition, and a restriction preventing consequential use before completion. “To be agreed” is not an adequate description of an essential safeguard on which the proposed deployment already depends.

The contract review should also identify enforcement limits. Relevant questions include the competent signatories, the order of precedence among documents, governing law, dispute forum, availability of interim measures, rights against upstream providers, and the effect of mandatory public or professional obligations. A contractual safeguard may be commercially negotiated yet legally ineffective, or legally enforceable yet too expensive to invoke. Both limitations belong in the approval record. No acceptance of a risk should be inferred merely from the absence of a recorded objection.

Evidence for rejection, restriction, or exit should be linked to a specific requirement. Examples include a material discrepancy in supplier disclosures, failure of an agreed local test, inability to obtain required records, unauthorised reuse, an unapproved dependency change, or failure to maintain a necessary exception route. These are proposed contractual triggers, not automatic statutory remedies. The institution must use the applicable procedure and preserve the evidence supporting its action. The resulting procurement file should show not only what was promised but why the institution believed the service could be supervised, corrected, and replaced.

5.3 Impact Assessment and Local Validation

An assessment of a proposed use, not a general certificate

An impact assessment should describe the change that an institution proposes to make and the evidence supporting that change. It is not a certificate attached permanently to a model. The same model can be used to locate public documents, prepare an internal summary, or generate material that an applicant is expected to treat as authoritative. Each use has different participants, information, consequences, and opportunities for correction. The assessment should therefore identify the system version and the institutional workflow together.

The first question concerns the legal objective. An institution should state the function it is authorised to perform, the existing difficulty, and the proposed contribution of the system. The second concerns affected interests: participation, confidentiality, equal treatment, timely consideration, professional advice, or another interest relevant to the particular setting. The third concerns mechanisms of harm. An assessment should

explain how a misleading translation, omitted record, incorrect category, inaccessible interface, or unreviewed recommendation could affect a person. Listing "bias" or "privacy" without this connection does not establish what must be tested or changed.

Statutory assessment requirements must be identified separately from the institution's proposed methodology. Section 31 of Kenya's Data Protection Act is one such statutory reference for data-protection impact assessment; applicable content, consultation, submission, and timing requirements need to be checked for the proposed processing. A general institutional form does not replace those requirements (Data Protection Act 24 of 2019, s. 31). Nor does completing a privacy assessment necessarily answer the different questions of substantive authority, legal accuracy, procedural access, or professional responsibility.

Pilot evaluation answers a bounded question about a specified workflow. The design must identify which uncertainty the trial may resolve, which safeguards must already work, and which possible consequences make experimentation unacceptable. A preliminary study can examine whether staff can retrieve and check a source without demonstrating durable improvement in access to justice. The worked resource example in section 5.4 makes this distinction explicit by separating assumed test results from the proposed decision they would justify.

Baselines and realistic alternatives

A baseline should describe the actual process, including both its deficiencies and the work that protects people. The institution should record the route into the service, the information requested, the time spent by staff and users, the treatment of incomplete matters, the opportunity to explain exceptional circumstances, and the route to review. Existing errors should be investigated rather than assumed away. However, a weakness in current practice does not establish that a particular AI proposal will address it.

The comparison needs a common unit of outcome. Documents processed, applications completed, matters resolved, and errors corrected are different measures. An intake system could increase the number of classified files while leaving the wait for urgent consideration unchanged. A translation assistant could shorten drafting time while increasing the work required to verify material legal terms. The assessment should connect intermediate activity to the person's legal objective and preserve the time and burden at every consequential stage.

A feasible non-AI alternative should be assessed against the same problem. For the hypothetical registry, alternatives might include a revised urgency form, clearer instructions, a staffed screening desk, a searchable index, or a simple rule directing all submissions containing a stated urgent request to an officer. These are not presumed superior. Their cost, error, accessibility, and supervisory requirements should be ex-

amined alongside those of the proposed classifier. The comparison is between achievable arrangements, not between an untested model and an idealized manual service.

Baseline weaknesses should also affect the permitted conclusion. Where the institution has no dependable record of abandoned applications or correction times, the assessment cannot demonstrate an improvement in those outcomes. It may first need a period of observation or records repair. That preliminary work can itself clarify whether AI is relevant. Missing baseline values should remain visibly missing; they should not be replaced by a convenient estimate presented as observed practice.

Local testing and the meaning of an error

Testing should begin by identifying the output that matters and the consequence of getting it wrong. Search requires attention to relevant omissions as well as irrelevant results. Translation requires attention to obligations, qualifications, deadlines, names, and remedies rather than fluency alone. Classification requires attention to which wrong categories change access or delay. Generated text requires verification against the record and applicable sources. These are proposed task distinctions; the acceptance criteria must be specified by the institution with suitable legal and technical input.

A test collection should cover ordinary work, difficult cases, and known failure conditions. The institution should identify how cases were selected, which populations and languages they represent, and which remain outside the evidence. It should record the relationship between development and evaluation material so that a supplier's prior exposure does not silently undermine the comparison. A test on curated documents should not be described as evidence about poor scans, handwritten forms, mixed-language submissions, or interrupted sessions unless those conditions were actually examined.

Legal evaluation also requires a procedure for disagreement. A fabricated citation is different from a contested interpretation of a valid authority. A disputed urgency classification may reflect uncertainty in the source narrative rather than a simple labelling mistake. Reviewers should distinguish factual error, legal disagreement, inadequate evidence, and matters outside the permitted scope. Where the benchmark answer is contested, the record should preserve the reasons and the decision about how that case affects acceptance.

The voluntary NIST AI Risk Management Framework places attention on context, measurement, and documented decisions about whether a system should proceed, and includes consideration of non-AI alternatives and residual risks. It is used here as a technical reference, not as law in the focal jurisdictions (National Institute of Standards and Technology [NIST], 2023, MAP, MEASURE, and MANAGE 1-2). The proposed legal assessment adds the specific authority, affected-person procedure, and remedial responsibilities relevant to the institutional task.

Table 5.2. Task-specific local validation record

Task	Material failure to test	Evidence required	Acceptance question to resolve in advance
Legal search	Missing current or materially adverse authority; wrong jurisdiction	Defined test questions, relevant source set, omission review and currency check	Can the intended user locate and verify the material necessary for the task?
Translation	Changed deadline, remedy, obligation, qualification or identity	Source and translated versions; qualified language and legal review; user comprehension exercises	Is legally consequential meaning preserved in the languages and conditions covered?
Intake or urgency classification	An urgent matter becomes invisible or is routed to an unavailable queue	Category definitions, challenging examples, queue histories and exception exercises	Can staff identify and correct consequential misrouting before avoidable harm?
Drafting or summarisation	Unsupported proposition, invented source or omitted material fact	Source-linked review of propositions and omissions; actual checking time	Can a responsible reviewer verify the work within the funded workflow?
Record matching or extraction	Wrong person, missing page, altered number or unreported uncertainty	Original records, transformation history and difficult-case testing	Is uncertainty visible, and is a non-adverse route available for unresolved matches?
Human and service controls	Reviewer cannot intervene; fallback unavailable; complaint loses records	Observed exercises under realistic workload, interruption and absence conditions	Do the proposed safeguards function without exceptional supplier intervention?

The table deliberately supplies no universal pass percentage. The institution must state its chosen criterion, why that criterion fits the consequence, the evidence supporting it, and the uncertainty remaining after testing. Absence of an observed error in a limited collection does not establish that the error cannot occur. An acceptance record should explain which risks the test could detect, which it could not, and how the permitted scope reflects that distinction.

Adversarial testing and confidential legal material

A local test must include deliberate interference where the proposed workflow makes it plausible. NIST's generative AI profile identifies prompt injection and data poisoning as information-security risks, and its adversarial-machine-learning taxonomy examines poisoning, indirect prompt injection and related privacy compromise. These technical accounts supply categories for constructing tests; they do not establish that a particular legal institution has experienced an attack or that adopting a named defence guarantees safety (NIST, 2024, §2.9; Vassilev et al., 2025, §§3.2, 3.4).

The threat model starts with trust boundaries. In a legal-research service, a retrieved judgment, party submission or web page is material to be analysed; it is not authorised to change the institution's instructions. In the registry example, an attachment may ask the system to hide the application, reveal another file or contact an external address. An instruction appearing inside that attachment remains untrusted even when it resembles an official system message. Testing must determine whether the application separates document content from its operating instructions and whether a compromised output can reach a consequential action. A model's refusal in a demonstration is weaker evidence than a boundary enforced by the surrounding software.

The institution should therefore begin with the least access required for the approved function. A search or translation assistant ordinarily has no need to send communications, alter a case record or retrieve every client file. If a connected tool is needed, its permissions, permitted recipients and human approval point should be explicit. The test should attempt to cross each boundary using synthetic documents and fictitious identifiers before confidential material is introduced. Such exercises must be authorised and contained; an evaluator should not test leakage by uploading actual client secrets to an unapproved service.

Poisoning presents a different failure mechanism. A retrieval collection can acquire an altered authority, a fabricated document or misleading metadata even though the foundation model remains unchanged. The relevant controls include admission of documents into the collection, source checks, version history, separation of draft and operative materials, and a way to withdraw a contaminated item. The test should ask whether an unauthorised replacement can change the answer unnoticed and whether earlier answers that relied on the item remain traceable. Replacing the model alone would not repair an untrustworthy source collection.

Confidentiality testing should examine both access and output. The institution can plant harmless canary records containing unique fictitious strings in separated test accounts and ask whether a user can obtain material beyond the permitted matter. It can also test whether logs, troubleshooting exports or support interfaces expose those strings. A clean result supports only the paths, settings and attack attempts actually tested. It cannot establish that all possible extraction attacks have been excluded, es-

pecially after a provider change. The acceptance record should state this residual uncertainty and the operational restrictions used to limit its consequences.

A useful attack record preserves the test document, system configuration, authorised permissions, attempted boundary crossing, observed behaviour and reviewer's decision. Failures should be classified by consequence. A misleading answer confined to an isolated test environment is different from disclosure across client matters or an unapproved external action. Any tested path that permits confidential cross-matter disclosure or an unauthorised operative action should block that capability until containment and retesting are complete. This is the chapter's proposed acceptance rule for the stated workflow, not a numerical legal standard. Permitting a narrower, read-only configuration requires evidence that the dangerous path is actually unavailable.

These tests also change procurement drafting. The Test Schedule should name the adversarial scenarios and the evidence the supplier must provide; the Dependency Schedule should identify who can alter retrieval, permissions and upstream services. The Incident Schedule should include successful boundary crossing and credible near misses, even when ordinary performance measures remain high. Supplier cooperation must cover relevant logs, affected accounts and changed components without requiring the institution to prove final causation before containment. Testing costs, remediation responsibility and repeat assessment after material changes must be priced alongside routine accuracy evaluation.

Human review remains useful but cannot carry the entire security burden. An officer may notice a false legal proposition while being unable to detect that the same interaction transmitted confidential data elsewhere. Similarly, reviewing the final letter does not undo a system's earlier alteration of a filing status. The assessment must locate each irreversible or externally consequential step and put the relevant control before it. Where that cannot be demonstrated, the supported decision is to remove the capability, confine the use to non-confidential material, or decline deployment rather than approve a broad service subject only to a warning.

Language, equality, and affected-person participation

Local validation concerns the whole service rather than the interface alone. A person may submit a narrative in one language but receive a legally consequential response in another. An accessible initial form may lead to a review procedure requiring unsupported documents or an unavailable communication channel. The assessment should follow a matter through entry, explanation, notification, correction, and review. It should identify where linguistic or accessibility support is needed and who can provide it when the automatic function is uncertain.

Group-specific evaluation should follow a reasoned account of foreseeable differences, not a list copied from a generic framework. Relevant questions may concern language, region, disability, age, documentation, or a combination of circumstances. The lawful

basis and necessity for collecting any sensitive evaluation data must be considered. Small or incomplete groups should not be reported as confidently equivalent merely because the overall result appears acceptable. The alternative may be a narrower deployment, additional evidence, or a supported route outside the unvalidated function.

Participation can identify difficulties that designers and legal reviewers have overlooked. It should involve an understandable account of the proposed task, what participants can influence, how their contributions will be used, and whether participation affects access to services. Individuals should not need to disclose the substance of a sensitive dispute merely to comment on a form or notification. The assessment should preserve disagreements and the institution's response, rather than treat attendance or an absence of objections as endorsement.

Independent scrutiny should be matched to consequence and available evidence. For a bounded internal research tool, a reviewer outside the immediate delivery team may be able to examine the test and its limitations. For a system influencing access to urgent relief or a large population's treatment, greater separation, specialist input, and secure evidence access may be necessary. Independence is not established by an external invoice alone: the reviewer's role, conflicts, access, and freedom to report adverse findings should be documented.

Why warnings and nominal human review can be inadequate

A warning cannot supply missing legal authority or make an unauthorised disclosure reversible. Nor can a disclaimer repair a service that prevents a person from submitting information before any reviewer sees it. The institution should identify the point at which a consequence occurs. Where harm precedes review, a promise that an official will check the eventual output does not address that harm. The appropriate response may require limiting collection, changing the workflow, removing an automated gate, or postponing live use.

Human review has similar conditions. The reviewer needs the original record, an intelligible account of the proposed output, competence in the task, time to investigate, and authority to depart. In the hypothetical registry, requiring approval of every suggested queue is not meaningful if the officer sees only the category and cannot open the submission. A test should therefore examine whether the reviewer detects deliberately introduced errors and can route a case outside the ordinary categories. The requirement is demonstrated capability, not the presence of a signature field.

Some limitations can be managed through narrower scope. A translation tool might be restricted to preliminary internal orientation while legally operative notices receive qualified review. A classifier might suggest categories while all submissions remain visible in a common queue. Other limitations may be incompatible with the proposed service, such as an inability to prevent prohibited secondary use or to preserve evid-

ence essential to review. The assessment should distinguish redesignable conditions from defects that remain unresolved under the intended configuration.

Authorizing and renewing a limited pilot

A pilot should have a defined question, permitted sites and users, data basis, duration or review event, monitoring responsibilities, and stopping conditions. The decision must distinguish an isolated test, shadow evaluation, and a live service. Shadow operation limits reliance on outputs but can still involve real personal or confidential information. A live pilot that affects a person's route through a legal process is not rendered legally inconsequential by its experimental label.

The evidence sufficient for a limited pilot is therefore bounded. It should establish the lawfulness of the trial activities, the feasibility of the proposed controls, a plausible connection between the task and the identified need, and an ability to detect and contain the foreseeable failures relevant to the trial. It need not pretend to prove durable effectiveness before observation is possible. However, any uncertainty left for the pilot to investigate must be compatible with the consequences to which participants will be exposed.

Table 5.3. Impact-assessment and pilot-authorisation template

Assessment field	Required record	Responsible approval or scrutiny
Purpose and authority	Defined problem, legal function, jurisdiction, relevant provisions and excluded uses	Competent institutional authority with legal review
System and workflow	Version, inputs, outputs, users, decision points and dependencies	Operational owner and technical reviewer
Affected persons and rights	Population, access routes, potentially affected interests and foreseeable harm mechanisms	Legal, domain, language and accessibility contributors
Baseline and alternatives	Existing process, known errors and burdens, missing evidence, feasible non-AI comparison	Evaluation lead and operational owner
Data and confidentiality	Sources, permissions, sensitive material, retention, access and secondary-use restrictions	Data steward with legal and security review
Validation	Test design, material failures, results, uncertainty, excluded settings and acceptance criteria	Qualified evaluators with specified independence

Table 5.3. Impact-assessment and pilot-authorisation template

Assessment field	Required record	Responsible approval or scrutiny
Human control and support	Demonstrated review, exception handling, assistance and fallback under realistic demand	Supervising officer and service-support lead
Participation	Who was consulted, access arrangements, disagreements, responses and unrepresented interests	Assessment owner; independent scrutiny as appropriate
Residual risks	Unresolved issues, their consequences, proposed restrictions and conditions not capable of being met	Competent authority; legal escalation where necessary
Pilot decision	Permitted scope, users, dates or events, monitoring, stop authority and resumption conditions	Named institutional signatory and designated alternate
Material change	Change description, affected assumptions, repeat tests and revised conditions	Change authority with relevant specialist review

The approving record should state whether the activity is permitted within a defined scope, requires redesign, awaits additional evidence, or cannot proceed as proposed. Conditions should be attached to an owner and verifiable completion event. The assessment should be renewed when a material change affects its assumptions, including changes in legal criteria, model or retrieval sources, working languages, population, data quality, staffing, or the role of human review. A previously completed form should never be treated as approval of an undocumented later use.

5.4 Staffing, Training, and Financial Planning

Staffing as the capacity to perform specified work

The minimum credible staffing arrangement cannot be determined from the word "AI." It depends on the task, volume, consequences of error, service hours, available expertise, and ability to use an alternative process. An internal tool that helps a small team search public sources creates a different pattern of work from an intake service used by people seeking urgent assistance. The staffing analysis should identify what must be done, when it must be done, who can do it, and who acts when the usual person is absent.

The implementation plan should distinguish responsibilities from job titles. A smaller institution may combine several functions in one post and obtain specialist assistance through a shared service. A larger institution may allocate them to separate teams. In either arrangement, someone must retain authority over the institutional use, someone must understand the data and records, and someone must be able to address an affected person's problem. Combining roles should not conceal a conflict between delivering the project and independently assessing it.

NIST's voluntary framework provides a useful technical reference for documenting roles, training responsibilities, leadership accountability, system inventories, and arrangements for decommissioning. These are governance categories, not prescribed headcounts or domestic public-service establishments (NIST, 2023, GOVERN 1.5-1.7 and 2.1-2.3). The institution must translate each required function into an actual assignment, with sufficient time, access, competence, and an identified alternate.

A role-based implementation arrangement

The senior institutional owner should be accountable for the permitted scope, resources, approval conditions, and response to material findings. That role should not absorb responsibilities legally assigned to a judge, professional, regulator, data custodian, or other office. The plan should distinguish accountability for the service from authority to decide a particular matter. A technology programme manager may arrange training and maintenance without having power to reconsider an applicant's entitlement.

Legal review should remain available after launch. New purposes, contract variations, complaints, and incidents can raise questions that the original assessment did not resolve. Data stewardship should cover source meaning, quality, permissions, corrections, recipients, and retention. Records management should ensure that relevant decisions and versions can be reconstructed. Security support should cover authorised access, vulnerability handling, investigation, recovery, and the protection of preserved evidence. These functions overlap but are not interchangeable.

Operational staff require a practical route to specialist assistance. A frontline officer should know how to handle an uncertain translation, preserve a disputed output, and send an urgent matter to a person authorised to act. They should not be expected to diagnose a model architecture before obtaining help. Conversely, a technical specialist should not be expected to decide a legal question merely because they can explain how the output was produced. The escalation record should preserve both the technical account and the competent legal or professional determination.

Table 5.4. Role and responsibility schedule for institutional implementation

Role to assign	Continuing responsibility	Evidence of capacity and required coverage
Senior institutional owner	Approve permitted scope, resources and action on material findings	Written delegation or authority; available alternate; recorded decisions
Legal reviewer	Examine authority, procedure, professional duties, contract and disclosure issues	Relevant expertise; access to governing sources; defined escalation availability
Operational supervisor	Monitor workflow, workload, exceptions and permitted human intervention	Demonstrated case exercises; authority to pause affected functions; coverage during absence
Data steward	Maintain source meaning, permissions, quality and correction links	Dataset inventory, provenance, recipient and correction records; custody arrangements
Records manager	Preserve and retrieve approvals, versions, decisions, complaints and holds	Retrieval test, retention schedule, access controls and verified exports
Security lead or shared security service	Handle access, vulnerabilities, containment, investigation and recovery	Tested response and recovery arrangements; contact and escalation obligations
Procurement and contract manager	Enforce deliverables, disclosures, changes, evidence rights and exit	Contract register, milestone checks, renewal dates and access to legal support
Evaluation lead	Maintain testing, baseline comparison, uncertainty and outcome analysis	Task-relevant methods; protected access; documented independence and conflicts
Language and accessibility support	Test consequential communication and assist excluded or uncertain cases	Qualified review and realistic user exercises; availability across required channels
Complaint and remedy coordinator	Acknowledge, preserve, route and track reconsideration and implementation	Case-tracking test; connection to competent decision-makers; urgent-contact coverage
Finance owner	Record lifecycle commitments, variation exposure and continuity resources	Cost register, funding authority, assumptions, renewal and transition provision

Table 5.4 identifies eleven functions rather than eleven mandatory posts. An implementation record must go further by naming the people, authority, time allocation, access and alternates that make the assignment real. The compact worksheet below can be repeated for each function. Where one person performs several roles, the independence and workload entries should explain how conflicts and absence are handled. An unfilled essential function restricts the approved service; describing it as a shared responsibility does not provide capacity.

An assignment worksheet and common role vocabulary

Assignment field	Required entry
Function and person	Function from Table 5.4; named office-holder and organisational location
Authority and limits	Delegation or governing instrument; decisions the person may and may not take
Capacity	Allocated hours, demonstrated competence and work displaced by the assignment
Records and access	Systems and records available; method of obtaining protected specialist access
Cover and escalation	Named alternate, absence procedure, urgent contact and unresolved-demand route
Independence	Conflicts, separation from delivery, protection for adverse findings and review recipient
Resources and verification	Funding source or shared-service agreement; date and result of the last practical exercise

The role vocabulary is consistent across the two chapters. The institutional owner authorises service scope and resources within delegated powers. The decision-maker applies the substantive legal function in an individual matter. The complaint and remedy coordinator maintains the chronology and referrals; this is the coordinating function called the remedial owner in the chapter 4 pathway, and it does not confer power to decide an entitlement. The records custodian preserves and supplies authorised evidence. The independent reviewer examines findings without owning delivery. The exit owner coordinates transition and surviving obligations. Several functions may be combined, but the person who promotes deployment should not be the sole arbiter of a contested claim that essential controls are adequate.

Different resource levels and shared expertise

A small institution considering internal assistance may use an existing supervisor and records officer, supported by periodic legal and security advice. Its scope may need to remain limited to public materials and non-operative outputs until broader safeguards are available. A medium-sized service may require dedicated operational supervision and a regular independent evaluation arrangement. A high-volume or consequential service may need continuous support, specialist language coverage, and protected investigation capacity. These are illustrative operating patterns, not descriptions of all institutions in any country or automatic thresholds for permission.

Shared public expertise can reduce duplication of specialist work. A common testing service might examine a model's behaviour on a defined task; a pooled security team might support incident response; a shared procurement unit might maintain clause libraries. The receiving institution still needs to know what was tested, which configuration was covered, and who will assist when local conditions differ. Shared provision should not remove the local officer's ability to reach the records or obtain an urgent response.

A shared-service agreement should specify response commitments, priorities during simultaneous incidents, confidentiality, access, conflicts, funding, and responsibility for implementing recommendations. Without these arrangements, apparent capacity may depend on informal assistance that disappears when demand increases. The plan should test the service under plausible competing demands rather than assume that a specialist can be available to every participating institution at once.

The counterargument is that detailed staffing requirements could make useful assistance unaffordable. That concern supports a narrower service and proportionate organisation, not counting unavailable safeguards as if they existed. An institution may choose to improve records or use ordinary search while arranging the expertise needed for a more consequential application. It should compare actual alternatives and their continuing work requirements, rather than presume that either a complex AI service or the existing process is necessarily the affordable option.

Demonstrating competence through practical exercises

Attendance demonstrates participation in training, not the ability to perform a task. Competence should be assessed using the records, interfaces, languages, and time conditions encountered in the intended workflow. An exercise should require the participant to explain what they checked, identify uncertainty, make or seek the appropriate decision, and preserve the record. The assessor should distinguish an error in technical operation from an error in legal judgment or escalation.

For legal research, an exercise can include a plausible but nonexistent citation, an outdated provision, or a source from the wrong jurisdiction. The participant should locate the actual authority and verify the proposition rather than ask the same system to con-

firm its own answer. For translation, the exercise can alter a qualification or deadline while preserving fluent prose. A trained participant should recognise when qualified language or legal review is necessary and should not present an unchecked rendering as authoritative.

For administration, an exercise can involve an urgent submission without the expected identifier, a misrouted file, or a complaint linked to an earlier version. The participant should preserve the person's opportunity to obtain consideration while using the authorised exception route. A records exercise should require retrieval of the input, configuration, output, human action, and subsequent correction for a selected transaction. A continuity exercise should require staff to maintain the essential function without the AI service and reconcile matters when ordinary operation resumes.

Supervisors need assessment as well. They should demonstrate that they can distinguish an isolated defect from a possible recurring problem, authorise an appropriate restriction within their powers, and obtain specialist assistance. Evaluators should be assessed on their ability to explain what a test does not establish. Procurement staff should be able to identify a promise unsupported by an upstream agreement, while finance staff should identify continuing obligations absent from a quotation. These exercises connect competence to assigned responsibility rather than to general familiarity with AI terminology.

Training records should state the task, exercise version, assessor, result, limitations, required remediation, and the conditions under which the person may operate independently. A failure should generate supervised practice or a restriction, not merely another attendance certificate. Reassessment should follow material changes in the system or role and should also address skill loss where ordinary work has become heavily dependent on automated assistance. These are proposed institutional controls, not a universal professional qualification regime.

Whole-life cost and resource assumptions

A financial plan should distinguish acquisition, implementation, operation, change, and withdrawal. The price of a licence or a quoted transaction does not establish the cost of a completed, reviewable legal service. The institution should identify the work required to prepare and correct records, test local performance, verify outputs, assist users, investigate complaints, preserve evidence, and maintain an alternative route. Costs borne by affected people, including repeated travel or paid assistance, should be recorded separately from institutional expenditure.

A cost estimate is an argument whose assumptions can be examined. Each amount should identify its quantity, unit price, period and evidential basis. Quoted licence charges can be treated as contractual inputs, whereas future complaint rates and checking times may remain uncertain. Those uncertainties belong in alternative scenarios. The worked example below uses invented planning assumptions solely to show

the arithmetic and the decision consequences; it does not attribute prices, staffing or performance to an existing institution.

Table 5.5. Lifecycle cost and dependency register

Cost category	Inputs required before an estimate	Continuing dependency to test
Preparation and integration	Record quality, interfaces, migration work, permissions and staff time	Whether source systems and custodians can maintain the required data
Licensing and hosting	Unit price, usage basis, minimum commitments, taxes and applicable currency	Exposure to changing volume, terms, exchange rates or provider availability
Verification and supervision	Observed checking time, task mix, workload and escalation frequency	Whether savings remain after legally and professionally necessary review
Language and accessibility	Required languages, channels, qualified review and assisted-service demand	Availability during ordinary work, urgent matters and staff absence
Security and maintenance	Testing scope, support, updates, monitoring and recovery	Whether essential controls remain funded after initial acceptance
Evaluation and records	Baseline work, repeat testing, storage, retrieval and independent scrutiny	Capacity to investigate changes and reconstruct disputed decisions
Complaints and correction	Expected work derived from observed service evidence; case-review and propagation tasks	Availability of competent review and implementation support
Continuity and exit	Parallel service, export, transition, archives and disposal	Usable replacement or manual route; obligations surviving the subscription

The finance owner should separate committed expenditure from contingent exposure and distinguish known quantities from assumptions. Where demand is uncertain, the plan can use scenarios with clearly stated inputs rather than report one number as a reliable forecast. A claimed saving should account for verification, corrections, and additional support, and should identify whether the effect is cash released, time redistributed, or capacity to handle more work. These outcomes are not interchangeable.

Approval also requires attention to the authority to incur future expenditure. Multi-year licences, usage commitments, transition assistance, and obligations after ter-

mination may require different financial treatment under the applicable rules. The chapter does not assume a common budget regime across jurisdictions. The implementation record should name the competent financial approver, the funding period, restrictions on commitments, and what will happen if renewal is unavailable. Essential-service continuity should not depend on an unrecorded assumption that another budget will cover it.

A bounded resource and acceptance example

Consider a second constructed application of Case M: a small registry wants an assistant to suggest routing categories and help staff locate documents submitted with a complaint about a suspended benefit. The registry cannot determine the underlying entitlement. Its legal office has established authority for staff-assisted intake and a lawful basis for the proposed processing, but has not authorised automated rejection, external messaging or unrestricted submission of confidential documents to a public service. These are stipulated conditions for examining capacity, not a report that any named registry has obtained approval.

The planning unit is one month of 400 submissions across twenty working days. Two existing clerks each have twenty hours per month that management can release for the proposed work without reducing essential counter service. An operational supervisor can allocate eight hours, a records officer six, a shared security service four, and a legal reviewer four. A separate evaluator has an agreed six-hour allocation. Language assistance has a twelve-hour allocation covering the selected trial languages. Every quantity in this example is an invented planning input, not a staffing recommendation or observed workload. The purpose is to make a capacity constraint visible enough to change the decision.

Suppose the proposed checking protocol initially takes six minutes per submission. At 400 submissions, checking alone consumes forty hours and exhausts the clerks' combined allocation. If ten percent of submissions require a further fifteen minutes of clarification, an additional ten hours is needed. The workload is already infeasible before allowing for absence or complaints. Reducing the check to four minutes merely to fit the spreadsheet would require evidence that necessary work still occurs. The immediate alternatives are to narrow the trial volume, allocate additional supported time, or improve the intake process without the assistant.

The narrowed option permits a maximum of 200 submissions in the month and retains every submission in a visible common queue. Checking consumes twenty hours; twenty additional fifteen-minute clarifications consume five more. The remaining fifteen clerk-hours provide planned capacity for record completion, complaints and cover. This is a planning margin rather than proof that demand will fit. A weekly workload review that shows checking and clarification exhausting the remaining allocation triggers a pause in the assistant's trial intake; the ordinary assisted route continues. The cap is attached to human work, not merely the supplier's technical throughput.

For illustration, express costs in notional units, abbreviated NU, so that invented prices cannot be mistaken for a country's wage rates or market quotations. Value clerk time at 20 NU per hour, supervision at 35, records work at 25, security and legal review at 50, independent evaluation at 45, and language support at 30. These values allow comparison of resources committed to the trial; they are not necessarily additional cash payments. Existing salaries count as opportunity cost even where no new post is created. Procurement must separately identify which commitments require new spending authority.

Monthly resource assumption	Calculation in NU	Resource cost
Clerk allocation including cover	40 hours × 20	800
Operational supervision	8 × 35	280
Records work	6 × 25	150
Shared security	4 × 50	200
Legal review	4 × 50	200
Independent evaluation	6 × 45	270
Language support	12 × 30	360
Restricted hosting and service	Assumed quotation	300
Protected export and storage	Assumed quotation	80
Operating resource total	Sum of stated entries	2640
Illustrative twenty percent contingency	2640 × 0.20	528
Monthly planning provision	2640 + 528	3168

Assume a separate one-time preparation allocation of 900 NU for forms, test material, training and the first retrieval exercise. One trial month therefore requires a planning provision of 4068 NU, combining preparation and the monthly amount. The twenty percent contingency is an illustrative arithmetic choice, not evidence of the likely in-

cident cost or an adequate reserve for litigation. An institution replacing these inputs with real data should separate committed cash, staff time, contingent exposure and costs borne by users. It should also show what the ordinary service costs over the same period; this example cannot establish savings without that comparator.

Independence has an operational price. In this example the evaluator belongs to a shared public service outside the delivery team, receives a fixed allocation unrelated to approval, can inspect protected test records and reports directly to the approving office. The six hours are reserved in writing, and a named alternate covers absence. A supplier may explain a technical finding but cannot veto its inclusion. If those arrangements are unavailable, an external title on the assessment does not solve the problem: the registry must obtain a credible review arrangement, further restrict the experiment to isolated material, or postpone consequential use.

Acceptance is decided against conditions declared before testing. Stipulate an isolated collection of 120 constructed submissions, including forty deliberately difficult urgent or ambiguous matters and material in the two proposed trial languages. The numbers describe an exercise design, not a representative sample of the service population. The required gates are that all submissions remain retrievable, every explicit request for urgent human attention reaches the common queue, restricted records cannot be retrieved across matters through the tested paths, and staff can reconstruct and correct a disputed routing suggestion. Checking time must also remain compatible with the trial cap. Accuracy averaged across easy files cannot offset failure of one of these essential controls.

Now consider stipulated results for Decision A: the classifier makes eight wrong suggestions, all visible and corrected in the exercise; no submission is hidden; one hostile attachment causes an attempted external request that the application blocks; the source, output and intervention for every disputed example can be retrieved. The legal and technical reviewers still disagree about three urgency labels. The supported decision is permission for the restricted shadow trial only: outputs cannot control queue visibility or delay, the disagreement is retained, and the trial tests checking time and staff use of exceptions. No claim of improved legal outcomes follows. The blocked request remains a near miss requiring examination, not a success erased from the record.

For Decision B, change one result: the same hostile attachment causes material from another test matter to appear in the answer, or the classifier hides one explicit urgent request. The ordinary accuracy figure could be identical, yet the essential confidentiality or access condition has failed. The consequential trial is refused. A later read-only experiment using public or synthetic material may be considered only if its boundaries are independently demonstrated; a general promise of human review does not address harm that occurs before that review. The contrast shows why the acceptance method needs failure consequences and permissions, not a universal pass percentage.

Even Decision A leaves uncertainty. Zero observed queue-loss events in a small designed collection does not prove a negligible rate in live work; the examples were intentionally selected and cannot support a population estimate. The trial therefore preserves all submissions, watches actual workload, and limits reliance. If a later evaluation seeks to estimate an error rate, it needs a defined sampling frame, enough observations for the intended precision, treatment of repeated or correlated cases, and uncertainty reporting. The initial exercise answers whether specified controls can work under stated conditions, not how often the service will harm a population.

Case M then connects the procurement and remedy records. The complaint coordinator assigns one matter identifier to the incoming notice, identity document, routing suggestion, staff correction and transmission to the competent benefit authority. The registry does not declare the benefit restored when it forwards the matter. It records whether the receiving authority accepted the referral and whether the applicant was informed of the next step. A supplier replacement must preserve that chronology. The resource example thus treats successful routing, lawful substantive reconsideration and implemented relief as distinct outcomes with different responsible actors.

Phased implementation and documentary milestones

Milestones should depend on evidence rather than the arrival of a promotional launch date. A preparatory stage can improve records and define the task without exposing people to a new decision process. Procurement can follow once requirements and resources are sufficiently specified. Isolated testing, a bounded pilot, and routine operation should have distinct approvals. A later expansion should not be treated as merely another release where it changes the affected population or legal consequence.

Table 5.6. Phased implementation plan

Phase	Work and dependencies	Evidence required to move forward
Define the need	Map the process, authority, affected persons, baseline and alternatives	Approved problem statement; unresolved legal and evidential issues recorded
Establish capacity	Assign roles, shared support, funding assumptions and complaint routes	Named owners and alternates; credible resource and access arrangements
Specify and acquire	Complete requirements, due diligence, testing, evidence and exit schedules	Lawful procurement approval; deliverables and material uncertainties documented
Test without operative reliance	Use the approved data and environment; examine difficult cases and controls	Version-specific results; remediation; demonstrated retrieval and fallback

Table 5.6. Phased implementation plan

Phase	Work and dependencies	Evidence required to move forward
Conduct a bounded pilot	Apply limited permission, monitor actual work, support participants and preserve records	Findings against predeclared criteria; independent scrutiny appropriate to consequence
Authorize routine use or redesign	Decide scope, renewal conditions, monitoring and ongoing resources	Reasoned approval, restriction or refusal; no unresolved essential condition assumed satisfied
Reassess and retire where appropriate	Review changes, incidents, results, costs and continuing need	Updated decision; continuity, evidence and transition verified

The implementation plan should be maintained jointly by operational, legal, procurement, and financial owners. It should identify dependencies that can prevent a milestone from being completed, such as a missing data agreement, an unavailable translator, or an export that cannot be interpreted. Delays should be explained in relation to these conditions, not concealed through a declaration that the system is technically ready. Readiness concerns the service that the institution can responsibly operate, including the work required when the system is wrong.

5.5 Evaluation, Incident Response, and System Withdrawal

Monitoring the legal service after adoption

A decision to deploy establishes a permitted use under stated conditions; it does not settle whether those conditions will continue to hold. Source records change, staff leave, workloads fluctuate, legal criteria develop, and suppliers modify services. Evaluation must therefore examine the continuing relationship between the system and the institution's task. The relevant question is not simply whether the application remains available, but whether people can still obtain lawful consideration, intelligible communication, correction, and an effective response to a challenge.

Monitoring should preserve the difference between technical activity and legal outcomes. The number of generated documents does not establish the number of adequately verified documents. A completed referral does not establish that the receiving service accepted the case. Closure of a complaint does not establish that the disputed decision was reconsidered or that relief was implemented. Indicators should identify the event measured, the relevant population, the observation period, and the excluded or missing cases.

The proposed monitoring record should include errors by task and consequence, delay at consequential stages, use of exceptions, corrections and reversals, complaints, support needs, availability, and the full cost of operating the safeguards. Where justified and lawful, it should examine differences across relevant languages, locations, channels, or groups. An observed difference is a reason for investigation; it is not by itself a legally established finding of discrimination. The analysis should examine case mix, missing information, institutional rules, and other explanations without allowing those possibilities to become reasons for indefinite inaction.

User and staff reports should supplement, not replace, the operational record. Low complaint numbers can reflect satisfactory service, but they can also reflect an inaccessible complaints route or uncertainty about what happened. High use of an exception route may reveal poor classification, an unsuitable standard form, or a successful channel for cases that require individual attention. The interpretation requires evidence about the workflow. A dashboard should make those uncertainties visible rather than turn each number into an automatic verdict.

NIST's framework includes post-deployment monitoring, user input and appeals, incident response, recovery, change management, and decommissioning within its management activities. These provide technical reference points, not a substitute for jurisdiction-specific notice and remedial duties (NIST, 2023, MANAGE 4.1). The institutional record proposed here connects each finding to the official who can change the process and the person responsible for addressing already affected cases.

A review calendar with event-based escalation

A calendar gives routine review a predictable place in institutional work. It should not postpone action on a serious finding until the next meeting. The proposed frequencies below are planning defaults for adaptation, not statutory deadlines. The institution should shorten or otherwise adjust them where the legal consequence, scale, rate of change, or evidence requires it. An infrequently used internal aid and a continuously operating consequential service need different observation arrangements.

Table 5.7. Model review calendar

Review point	Material to examine	Owner and resulting record
Before first use and every material release	Authority, configuration, permissions, validation, training, evidence access and fallback	Competent approving authority; dated release decision with permitted scope
During each operating session or service shift, where relevant	Availability, urgent exceptions, failed transfers and unresolved protective actions	Operational supervisor; handover and escalation record

Table 5.7. Model review calendar

Review point	Material to examine	Owner and resulting record
Weekly during a limited pilot	Material errors, checking time, complaints, accessibility, exceptions and trial boundaries	Pilot owner and evaluators; findings and scope decision
Monthly during routine operation	Task outcomes, delay, correction completion, workload, incidents and cost changes	Service owner with specialist input; action register and overdue matters
Quarterly governance review	Continuing need, distributional findings, material changes, supplier compliance and unresolved risks	Institutional authority with independent scrutiny as appropriate; reasoned continuation review
Before budget or contract renewal	Whole-life cost, alternatives, dependency concentration, export and transition capability	Finance and procurement owners with operational approval
On a serious incident, legal change or loss of an essential safeguard	Immediate exposure, affected population, authority to restrict, evidence and continuity	Designated response authority; event-specific restriction and investigation decision
After withdrawal or replacement	Open cases, preserved records, access revocation, disposal, correction and remaining obligations	Exit owner and records custodian; verified closure or outstanding-action record

The review should retain adverse findings and disagreements, not only the final dashboard. Every action should have an owner, completion evidence, a review event, and an escalation route if it remains overdue. A recurring finding should be linked to earlier occurrences so that the institution can distinguish a new defect from an unresolved known condition. Changing a metric or its denominator should be documented; otherwise apparent improvement may reflect a change in measurement rather than service.

Incidents beyond cybersecurity

An AI-related incident may involve security, but it may also arise when a system operates exactly as configured. A lawful submission could disappear from an officer's view because a category closes the file. A translated notice could omit a qualification. An update could remove a relevant source from a search collection. These hypothetical events concern legal or service consequences even if no attacker entered the system and no server failed. The reporting definition should therefore include serious and repeated consequential failures, loss of an essential safeguard, unauthorised use, and relevant near misses.

NIST's incident-response publication treats preparation and improvement as parts of the wider cybersecurity risk-management process and relates detection, response and recovery to that framework. It is a technical security reference rather than an AI-specific legal-remedy protocol (Nelson et al., 2025, §2.1, pp. 4-6). The proposed institutional procedure extends the response record to the affected legal process: the transaction, person, decision, deadline, available interim action and responsible decision-maker. Technical recovery and legal correction should be coordinated but recorded separately.

Notification duties require their own legal analysis. For example, POPIA distinguishes an operator's immediate notification to the responsible party from the responsible party's notification obligations following a security compromise, with the statutory conditions and timing in section 22. A contractual reporting interval cannot replace those requirements (POPIA, ss. 21(2), 22). This illustration should not be converted into a common reporting deadline across jurisdictions or applied automatically to every non-security service error.

Initial reporting should not depend on a complete diagnosis. Staff should be able to report an unexplained consequential output, missing record, or failed exception route without first establishing the responsible component or legal breach. The first record should state what is known, what remains uncertain, which service is affected, and whether delay could aggravate harm. Allegations about supplier or staff responsibility should remain distinguishable from verified findings.

Table 5.8. Institutional incident-response protocol

Stage	Required action and evidence	Responsible role and protection
Receive and register	Assign an incident identifier; record time, reporter, affected service, known facts and uncertainty	Accessible reporting contact; urgent reports reach a person rather than an unattended queue
Assess immediate exposure	Identify continuing reliance, imminent deadlines, affected persons and essential services	Authorized operational and legal decision-makers; record any protective restriction
Preserve and contain	Secure relevant inputs, outputs, versions, human actions and access records; limit harmful use	Records and security leads; preservation does not authorise unrelated processing
Notify and coordinate	Identify applicable statutory, contractual and institutional notifications and competent recipients	Named legal and response contacts; do not delay required notice awaiting final causation

Table 5.8. Institutional incident-response protocol

Stage	Required action and evidence	Responsible role and protection
Investigate scope and cause	Examine source data, transformations, configuration, supplier changes and workflow; identify related cases	Qualified investigators with appropriate independence and protected access
Correct process and affected cases	Repair the relevant component, propagate corrections, support reconsideration and implementation of relief	Competent custodians and decision-makers; technical repair alone does not close the event
Decide resumption or exit	Review evidence, repeat necessary tests, verify resources and approve a bounded next step	Authorized signatory distinct from the supplier's declaration that the problem is fixed
Monitor completion	Check recurrence, unresolved harm, remaining records and implementation obligations	Service owner and independent reviewer as appropriate; dated completion evidence

Correction, restriction, suspension, and termination

These responses have different objects. Correction changes a record, configuration, procedure, or particular outcome. Restriction narrows the permitted use or population. Suspension temporarily stops a function or a form of reliance. Termination ends the contractual relationship or another defined arrangement. Withdrawal removes the system from the service. A decision record should identify which response is being taken and why; using the word "stop" without specifying the object can leave harmful processing active or unnecessarily interrupt an essential service.

An isolated and controllable formatting defect might be corrected while use continues within its approved scope. A repeated error in a consequential translation could require suspension of that language function while qualified translation continues. Evidence of prohibited secondary processing may require immediate containment and legal escalation even where outputs remain accurate. These are proposed response distinctions, not automatic legal conclusions. The competent authority must consider the applicable rules, available evidence, and consequences of both continued use and restriction.

Brazil's March 2025 Resolution 615 provides a bounded legal illustration. Its provisions distinguish corrective measures and suspension from discontinuation where discriminatory bias cannot be removed, with a documented justification. The rule concerns the judicial framework addressed by that instrument; it is not a universal termination clause for all AI services (CNJ, 2025, art. 8, §§2-3, English translation con-

sulted). The general analytical lesson is to specify the relationship between a finding, an available corrective measure, and the evidence needed to justify continuing reliance.

An institution should not require an affected person to prove an entire systemic mechanism before it considers urgent protection. It can assess the credible risk to the particular matter and use the lawful interim measures available to the competent body. At the same time, urgent restriction should preserve the evidence necessary to test the allegation. Protective action, investigation, findings of breach, and final liability are separate decisions. Their records should not be collapsed into a single incident label.

Withdrawal without loss of evidence or essential service

Withdrawal begins with identifying what depends on the system. The inventory should cover active matters, retained records, integrations, provider accounts, retrieval collections, configuration files, staff procedures, user notices, and complaint routes. It should distinguish evidence required to explain earlier actions from information needed to continue future work. Ending access to an interface must not silently remove the only way to interpret the institution's decision history.

The exit owner should verify that exported records are complete and intelligible for the intended purpose. A file count is insufficient where links between inputs, outputs, versions, and decisions have been lost. Where a proprietary component cannot be transferred, the record should state that limit and preserve the evidence available to investigate its earlier use. A replacement should be assessed in its own configuration; migration is not proof that prior validation applies to the new service.

Essential-service continuity may require temporary staffing, a manual route, a restricted replacement, or staged transition. The choice should be tied to the legal function and available resources. An institution should not keep a disputed automated gate in operation solely because it has not planned an alternative. Nor should it assume that abrupt shutdown is harmless where people have pending filings, assistance requests, or review deadlines. The transition record should identify those matters and the officers responsible for preserving their position.

Deletion should be controlled by the applicable retention and preservation requirements. A lawful investigation or litigation hold may require restricted preservation even after ordinary use ends. When disposal is due, the institution should record the material covered, method, responsible actor, exceptions, and verification. Closing the supplier's account does not necessarily delete all copies, revoke every integration, or resolve downstream decisions. Conversely, an unresolved claim is not a justification for indefinite unrestricted retention of all service data.

Resumption after suspension should be a new, reasoned decision. The institution should identify what was wrong, what changed, which tests were repeated, which affected matters were addressed, and what uncertainty remains. A supplier's statement

that the model has been improved cannot establish that the institution's procedure, staffing, or notice problem has been repaired. Where the necessary evidence cannot be obtained, a restricted alternative or continuing suspension may remain the only supportable operational position under the institution's own approval conditions.

Continuation-or-withdrawal decision form

Table 5.9 consolidates the record needed for a decision about continued use. It is not a universal numerical score or a substitute for the competent authority's legal judgment. Each finding should identify its source and the distinction between an observed fact, a disputed allegation, an estimate, and an unresolved question. The form should be signed by the authorised office-holder, with relevant specialist views retained even where they differ from the final decision.

Table 5.9. Continuation, restriction, suspension, or withdrawal decision form

Field	Entry required
System and decision	Institution, service, configuration, date, review trigger and proposed action
Competence	Decision-maker, relevant authority, scope of power and required procedure
Continuing need	Current problem, affected population, baseline and feasible alternatives
Findings	Performance, errors, delays, access, complaints, incidents, costs and evidence limitations
Legal and contractual conditions	Applicable requirements, unmet conditions, supplier obligations and unresolved interpretations
Affected persons	Known cases, potentially related cases, imminent harm and assistance arrangements
Available responses	Correction, restricted use, suspension, replacement or termination; consequences of each
Corrective work	Specific changes, owners, resources, evidence of completion and remaining uncertainty
Evidence and continuity	Preservation, export, pending matters, manual or replacement service, access and retention
Decision and reasons	Permitted scope or stopped function, conditions, responsible signatory and specialist disagreements

Table 5.9. Continuation, restriction, suspension, or withdrawal decision form

Field	Entry required
Follow-up	Required notifications, review event, resumption criteria, open remedies and closure evidence

Lasting improvement and the limits of comparative conclusions

Lasting improvement requires evidence that the institution can sustain a better legal service under ordinary conditions. Relevant evidence would connect reliable records and accountable decisions to timely consideration, usable communication, accessible review, and implemented correction. It would cover periods beyond a demonstration and explain changes in case mix, staffing, procedure, or demand that may also account for the result. This chapter does not report such a completed evaluation across the six jurisdictions.

Several elements of this handbook are transferable as questions and documentation methods. A defined task, an authority record, a source-and-version history, a tested exception route, an evidence-access agreement, a funded review function, and an exit inventory can organize inquiry in different settings. Their content is not automatically transferable. The body that may approve use, the lawful basis for processing, the procedure for disclosure, the remedy available, and the authority to suspend or terminate remain dependent on the relevant law and institutional arrangement.

Resource assumptions likewise cannot be transferred without examination. A shared testing service may make specialized expertise available to several institutions, but it does not automatically provide local language support or urgent reconsideration. A detailed audit clause may be usable where qualified reviewers can access and interpret the evidence yet remain ineffective where no one is funded to invoke it. Comparative analysis should therefore explain the conditions under which an arrangement could operate, rather than claim that a common form creates equivalent protection.

The final criterion is the institution's ability to answer a concrete sequence of questions. What was it authorised to do? What information and system configuration did it use? Who examined the output and made the consequential decision? What could the affected person understand and challenge? What was corrected, by whom, and with what result? Can the same service continue lawfully when the AI component is unavailable? These questions connect innovation to legal accountability without assuming that every application is transformative or that transformation is necessarily improvement.

Chapter 5 Key Arguments

Instrument choice and implementation must be assessed together. An institution needs a lawful function, usable evidence and people capable of acting on findings. Neither a detailed contract nor a completed assessment supplies those conditions by itself.

The worked registry example shows how workload, confidentiality and independence can change an approval decision. A technically promising service may justify only an isolated or restricted trial; an essential access or privacy failure can defeat approval even where aggregate task accuracy appears favourable.

Procurement and monitoring should preserve the relationship between the accepted configuration and the actual service. Changes in source collections, permissions, staff capacity or provider dependencies can invalidate earlier assumptions and require a new decision.

Implementation is complete only to the extent demonstrated by service outcomes and surviving responsibilities. Corrected data, operational recovery and closure of a supplier contract must be distinguished from lawful reconsideration and relief for people already affected.

Conclusion

What changes when computation reorganises legal work

The book began with a distinction between installing a system and transforming the operation of law. The comparisons support a more precise answer. Transformation occurs where computational arrangements materially redistribute control over a legally consequential opportunity, interpretation or corrective act. The decisive change may occur before a formal decision: in a referral, an eligibility flag, the sources visible to a lawyer or the records available to a reviewer. An unchanged statute and a human signature therefore do not establish unchanged legal practice.

This is a conditional account, not a claim that every AI application centralises power or diminishes rights. A searchable archive with effective source checking may improve existing work without materially changing legal relationships. A correction tool may strengthen an institution's ability to discover and repair errors. The transformation inquiry asks what changes and through which mechanism; the evaluative inquiry asks whether that change is lawful, justified and sustainable. Keeping these questions separate avoids treating innovation as either an inherent benefit or a sufficient ground for prohibition.

What the comparisons establish

The comparison of South African customary law and Kenyan religious jurisdiction establishes why legal representation cannot be reduced to database coverage. Shilubana requires attention to the development of customary law and constitutional norms; article 170 of Kenya's Constitution attaches jurisdiction to specified subject matter and party conditions. The constructed applications expose different errors: freezing a source of law and replacing jurisdictional facts with an inferred identity. The legal response must address the particular error. Adding more historical records or improving demographic prediction would not, by itself, resolve either problem.

Colombia and Brazil show how similar judicial tasks can sit within different governance instruments. A constitutional judgment, implementing judicial agreement and council resolution have different institutional origins and scopes. Their requirements concerning assistance, human responsibility, disclosure and record preservation must be read in their own terms. The completed comparison demonstrates that an abstract commitment to human control is insufficient to identify an institution's du-

ties. It does not establish how consistently courts follow those duties or what effect the instruments have had on case outcomes.

The data and remedies analyses establish a second distinction: technical correction, lawful reconsideration and relief are separate acts. A database operator may repair a false identity match while a benefit decision remains operative. A civil claim against a service provider may require proof of a different duty, causal connection and kind of loss from an administrative challenge. The South African and Brazilian worked applications therefore identify parties, governing provisions, possible defences, periods and forums before proposing a remedy. They do not treat every AI error as strict liability or every explanation request as a substitute for timely proceedings.

The evidence comparison establishes the importance of the proposition for which a record is offered. Preservation, authenticity, admissibility, expert assessment and weight answer different questions. India's statutory electronic-record requirements do not make the contents of every admitted record true. The Pakistani video-evidence discussion in Sarki exposes the need to examine judicial reasoning critically alongside the ultimate disposition. These analyses support a restrained conclusion: technical provenance assists proof, but its legal significance depends on the applicable rule, the contested proposition and the opportunity to challenge it.

These differences give the comparative orientation its value. The selected jurisdictions are not illustrations of one regional deficit. They reveal distinct arrangements for recognising law, distributing authority and making correction effective. Colonial recording practices, conditional recognition, language and supplier dependence matter when a demonstrated mechanism connects them to the application. The resulting questions travel beyond the selected settings, but the domestic answers cannot travel without renewed legal analysis.

Why safeguards fail and what implementation must change

A central counterargument is that well-designed process can legitimate a harmful system by making it appear reviewable. The book accepts that risk. Logs can document a wrong that no one has power to repair. Disclosure can shift investigative work to a person unable to bear it. Human approval can conceal automatic reliance. For that reason, the proposed controls are judged by the action they enable and the burden they impose, not by their mere presence in a contract or assessment. If the essential corrective capacity cannot be supplied, restriction or non-adoption is a reasoned outcome.

Institutional incentives explain why assigning responsibility on paper is insufficient. A delivery team rewarded for throughput may treat exceptions as failures of compliance. A supplier may offer attractive demonstrations while retaining exclusive control over diagnostic evidence. A manager may count staff savings without allocating time to re-

view. These are hypotheses about incentives, not empirical findings about every institution. The instruments respond by separating delivery from adverse evaluation, pricing review and exit, preserving access to relevant evidence, and recording disagreement at approval. Their effectiveness remains a matter for operational testing.

The worked capacity examples demonstrate how the framework changes a decision. Once checking, clarification, language support and independent review are counted, a seemingly efficient proposal may justify only a limited trial. Essential confidentiality or access failures defeat reliance even where average accuracy is unchanged. Conversely, a narrowly bounded tool can be acceptable without proving perfection if its errors remain visible, its role is lawful and effective human work can be sustained. Proportionality requires matching controls to consequence; it does not mean removing a safeguard until a project fits its budget.

The contribution and the next research questions

The scholarly contribution is the connection between operational transformation and the allocation of authority, evidence and correction. The book refines a threshold for identifying change, applies it within differentiated legal orders, and translates the analysis into linked institutional records. It builds on technological due process, sociotechnical critique and reviewability scholarship while refusing to treat documentation as an end in itself. The practical contribution is a method for reaching a reasoned decision about a particular task, including a decision to narrow, postpone or reject it.

The next research task is to examine whether these proposed connections survive ordinary work. Priority questions concern the treatment of exceptions under workload pressure, the propagation of corrections across organisations, the accessibility of complaint routes, and the independence of evaluation when a supplier controls technical evidence. A credible study would observe the same legal function over a defined period, include non-users and unsuccessful applicants, and record concurrent procedural and staffing changes. It would also examine cases where a simpler alternative produces better outcomes.

The conclusion established here is narrower than a forecast of legal systems transformed by AI. Legal institutions remain responsible for the consequences they attach to computational work. A defensible transformation makes that responsibility exercisable: the relevant actor can understand the basis of the action, obtain and assess its evidence, hear a challenge and implement an available correction. Where those connections fail, technical capability has outrun institutional justification. Where they hold, AI can be considered as one means of improving a legally defined service, subject to continuing evidence and review.

References

Cases

Constitutional Court of South Africa. (2008). Shilubana and Others v Nwamitwa [2008] ZACC 9. https://www3.saflii.org/za/cases/ZACC/2008/9.html

Corte Constitucional de Colombia, Sala Segunda de Revisión. (2024, August 2). Sentencia T-323/24 (Expediente T-9.301.656). https://www.corteconstitucional.gov.co/relatoria/2024/T-323-24.htm

Sindh High Court, Circuit Court Larkana. (2025, May 20). Shafi Mohammed Sarki v The State (Criminal Appeal No. D-44 of 2024). https://caselaw.shc.gov.pk/caselaw/view-file/MjY2OTc1Y2Ztcy1kYzgz

High Court of South Africa, Gauteng Division, Pretoria. (2017, May 3). Association of Test Publishers of South Africa v President of the Republic of South Africa and Others (89564/14) [2017] ZAGPPHC 144. https://www.saflii.org/za/cases/ZAGPPHC/2017/144.html

Legislation and institutional instruments

Autoridade Nacional de Proteção de Dados (Brazil). (2024, August 23). Resolução CD/ANPD nº 19, with correction published August 18, 2025. https://www.gov.br/anpd/pt-br/acesso-a-informacao/institucional/atos-normativos/regulamentacoes_anpd/resolucao-cd-anpd-no-19-de-23-de-agosto-de-2024

Brazil. (1985). Lei nº 7.347, de 24 de julho de 1985 [Public Civil Action Law]. https://www.planalto.gov.br/ccivil_03/leis/l7347compilada.htm

Brazil. (1990). Lei nº 8.078, de 11 de setembro de 1990 (Código de Defesa do Consumidor) [Consumer Defence Code]. https://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm

Brazil. (1998). Lei nº 9.610, de 19 de fevereiro de 1998 [Copyright Law]. https://www.planalto.gov.br/ccivil_03/leis/l9610.htm

Brazil. (2002). Lei nº 10.406, de 10 de janeiro de 2002 (Código Civil) [Civil Code]. https://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm

Brazil. (2015). Lei nº 13.105, de 16 de março de 2015 (Código de Processo Civil) [Code of Civil Procedure]. https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13105.htm

Brazil. (2018). Lei nº 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais) [General Personal Data Protection Law; LGPD]. https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709compilado.htm

Colombia. (1991). Decreto 2591 de 1991: Por el cual se reglamenta la acción de tutela consagrada en el artículo 86 de la Constitución Política. https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=5304

Consejo Superior de la Judicatura (Colombia). (2024, December 16). Acuerdo PCSJA24-12243. https://actosadministrativos.ramajudicial.gov.co/Home/GetFile?nombreArchivo=PCSJA24-12243.pdf

Conselho Federal da Ordem dos Advogados do Brasil. (2015, October 19). Código de Ética e Disciplina da OAB, approved by Resolução 02/2015. https://www.oab.org.br/util/print?numero=02%2F2015&origem=Resolu%C3%A7%C3%B5es&print=Legislacao

Conselho Nacional de Justiça (Brazil). (2025, March 11). Resolução nº 615/2025. Diário da Justiça Eletrônico do CNJ, 54/2025, March 14, 2-17, as amended by Resolução nº 674, March 25, 2026. https://atos.cnj.jus.br/atos/detalhar/6001

India. (2023). Bharatiya Sakshya Adhiniyam, 2023 (Act No. 47 of 2023). https://www.indiacode.nic.in/indiacode/bitstream/123456789/20063/1/aa202347.pdf

Kenya. (2010). Constitution of Kenya, 2010. https://www.parliament.go.ke/sites/default/files/2023-03/The_Constitution_of_Kenya_2010.pdf

Kenya. (2019). Data Protection Act 24 of 2019. https://www.odpc.go.ke/wp-content/uploads/2024/02/TheDataProtectionAct__No24of2019.pdf

Legal Practice Council (South Africa). (2019). Code of Conduct for All Legal Practitioners, Candidate Legal Practitioners and Juristic Entities. General Notice 168, Government Gazette 42337, March 29, corrected by Notice 198, Gazette 42364. https://lpc.org.za/wp-content/uploads/2020/10/CODE-OF-CONDUCT.pdf

Ministry of Home Affairs, India. (2024, February 23). Notification S.O. 849(E): Commencement of the Bharatiya Sakshya Adhiniyam on July 1, 2024. https://www.mha.gov.in/sites/default/files/BharatiyaSakshyaAdhiniyam_24022024.pdf

Pakistan. (2004). Public Procurement Rules, 2004 (S.R.O. 432(I)/2004). Public Procurement Regulatory Authority. Historical text. https://epms.ppra.gov.pk/public-media?file=L3BkZi9wcHJhLXJ1bGVzLnBkZg%3D%3D

Pakistan. (2010). Benazir Income Support Programme Act, 2010 (Act XVIII of 2010). Gazette of Pakistan, August 18. https://bisp.gov.pk/SiteImage/Downloads/BISP_ACT_2010.pdf

Pakistan. (2026). Public Procurement Rules, 2026 (S.R.O. 1658(I)/2026, September 28). Public Procurement Regulatory Authority, official online text. https://epms.ppra.gov.pk/public/procurement-rules

South Africa. (1969). Prescription Act 68 of 1969. https://www.justice.gov.za/legislation/acts/1969-068.pdf

South Africa. (1988). Law of Evidence Amendment Act 45 of 1988. https://www.justice.gov.za/legislation/acts/1988-045.pdf

South Africa. (1996). Constitution of the Republic of South Africa, 1996. https://www.justice.gov.za/constitution/chp02.html

South Africa. (1998). Employment Equity Act 55 of 1998. Government Gazette 19370, October 19. https://www.gov.za/sites/default/files/gcis_document/201409/a55-98ocr.pdf

South Africa. (2000). Promotion of Administrative Justice Act 3 of 2000. https://www.justice.gov.za/legislation/acts/2000-003.pdf

South Africa. (2002). Institution of Legal Proceedings against Certain Organs of State Act 40 of 2002. Government Gazette 24112, November 28. https://www.justice.gov.za/legislation/acts/2002-040.pdf

South Africa. (2005). National Credit Act 34 of 2005. Government Gazette 28619, March 15, 2006. https://www.gov.za/sites/default/files/gcis_document/201409/a34-050_1.pdf

South Africa. (2013). Employment Equity Amendment Act 47 of 2013. Government Gazette 37238, January 16, 2014. https://www.labour.gov.za/DocumentCenter/Acts/Employment%20Equity/Amendment/Employment%20Equity%20Amendment%20Act%2047%20of%202013.pdf

South Africa. (2013). Protection of Personal Information Act 4 of 2013. https://www.justice.gov.za/legislation/acts/2013-004.pdf

South Africa. (2014). National Credit Amendment Act 19 of 2014. Government Gazette 37665, May 19. https://www.gov.za/sites/default/files/gcis_document/201409/37665gon389.pdf

South Africa. (2022). Employment Equity Amendment Act 4 of 2022. Government Gazette 48418, April 14, 2023. https://www.labour.gov.za/DocumentCenter/Acts/Employment%20Equity/Amendment/Employment%20Equity%20Amendment%20Act%20No4%20of%202022_14%20April%202023.pdf

Scholarship

Biega, A. J. (2023). Data repurposing through compatibility: A computational perspective [Preprint, version 1]. arXiv:2309.00939. https://arxiv.org/abs/2309.00939v1

Citron, D. K. (2008). Technological due process. Washington University Law Review, 85, 1249-1313. https://journals.library.wustl.edu/lawreview/article/id/6697/

Cobbe, J., Lee, M. S. A., & Singh, J. (2021). Reviewable automated decision-making: A framework for accountable algorithmic systems [University repository version of the FAccT 2021 paper]. https://doi.org/10.1145/3442188.3445921

Haseeb, M., & Vyborny, K. (2020, October 22). Data, discretion and institutional capacity: Evidence from cash transfers in Pakistan [Working paper]. American Economic Association meeting version. https://www.aeaweb.org/conference/2021/preliminary/paper/d6667hir

Hoque, S. T. (2023). Neocolonial digitality: Analyzing digital legal databases using legal pluralism. Asian Journal of Law and Society, 10(3), 516-549. https://doi.org/10.1017/als.2023.9

Magesh, V., Surani, F., Dahl, M., Suzgun, M., Manning, C. D., & Ho, D. E. (2025). Hallucination-free? Assessing the reliability of leading AI legal research tools. Journal of Empirical Legal Studies, 22(2), 216-242. https://doi.org/10.1111/jels.12413

Selbst, A. D., boyd, d., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In Proceedings of the Conference on Fairness, Accountability, and Transparency (pp. 59-68). ACM. https://doi.org/10.1145/3287560.3287598

Technical policy and other materials

African Union Commission. (2024). Continental artificial intelligence strategy: Harnessing AI for Africa's development and prosperity. https://au.int/en/documents/20240809/continental-artificial-intelligence-strategy

Chandra, B., Dunietz, J., Roberts, K., Lee, Y., Fontana, P., & Awad, G. (2024). Reducing risks posed by synthetic content: An overview of technical approaches to digital content transparency (NIST AI 100-4). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-4

Clark, J., Haseeb, M., Jalal, A., Siddiqi, B., & Vyborny, K. (2022, March). Using biometrics to deliver cash payments to women: Early results from an impact evaluation in Pakistan. World Bank, ID4D Evidence Note. https://documents1.worldbank.org/curated/en/099155004142238180/pdf/P1763410d1e1af00108e170e5754d04fed9.pdf

Information Regulator South Africa. (n.d.). POPIA resources. Retrieved October 20, 2026. https://inforegulator.org.za/popia/

National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

National Institute of Standards and Technology. (2024). Artificial intelligence risk management framework: Generative artificial intelligence profile (NIST AI 600-1). https://doi.org/10.6028/NIST.AI.600-1

Nelson, A., Rekhi, S., Souppaya, M., & Scarfone, K. (2025). Incident response recommendations and considerations for cybersecurity risk management: A CSF 2.0 community profile (NIST Special Publication 800-61, Revision 3). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-61r3

Organisation for Economic Co-operation and Development. (2024). Explanatory memorandum on the updated OECD definition of an AI system (OECD Artificial Intelligence Papers, No. 8). OECD Publishing. https://doi.org/10.1787/623da898-en

Public Buyers Community. (2025, March 5). Updated EU AI model contractual clauses, with commentary. https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/updated-eu-ai-model-contractual-clauses

South African Government, Department of Employment and Labour. (2024, December 5). Minister Nomakhosazana Meth announces commencement of Employment Equity Amendment Act reducing regulatory burden for small employers [Media statement]. https://www.gov.za/news/media-statements/minister-nomakhosazana-meth-announces-commencement-employment-equity

UNESCO. (2025, December 3). Guidelines for the use of AI systems in courts and tribunals [Publication description]. https://www.unesco.org/en/articles/guidelines-use-ai-systems-courts-and-tribunals

Vassilev, A., Oprea, A., Fordyce, A., Anderson, H., Davies, X., & Hamin, M. (2025). Adversarial machine learning: A taxonomy and terminology of attacks and mitigations (NIST AI 100-2e2025). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-2e2025

World Wide Web Consortium. (2013, April 30). PROV-Overview: An overview of the PROV family of documents (P. Groth & L. Moreau, Eds.). W3C Working Group Note. https://www.w3.org/TR/prov-overview/