
FROM DIGITIZATION TO DIGITALIZATION: WHY THE INDIAN JUDICIARY MUST MOVE BEYOND PDF-CENTRIC ELECTRONIC FILING

Rudra Prasad Rimal, LLM (Criminal Law),
Faculty of Law, the ICFAI University, Gangtok, Sikkim

ABSTRACT

The Indian judiciary has made significant progress in adopting information and communication technologies through the e-Courts Mission Mode Project, electronic filing (e-Filing), virtual hearings, and the National Judicial Data Grid (NJDG). These initiatives have substantially reduced dependence on physical records and have improved accessibility and transparency within the judicial system. However, the underlying architecture of electronic filing remains predominantly document-centric, relying on Portable Document Format (PDF) files as the primary unit of judicial information. Although PDFs effectively replicate paper records in digital form, they fail to harness the transformative potential of digital technologies by embedding critical legal information within unstructured documents. This paper argues that the prevailing PDF-centric paradigm represents the final stage of digitization rather than the beginning of true digitalization. A judicial information system designed primarily around uploaded documents inevitably limits automation, interoperability, intelligent case management, legal analytics, and the responsible deployment of artificial intelligence. Instead, the paper proposes a transition towards a data-centric judicial architecture in which structured legal data becomes the principal repository of judicial information, while PDF documents function as supporting evidentiary material. Drawing upon principles of digital governance, information systems architecture, and judicial administration, the paper distinguishes between document management and information management within courts. It demonstrates that the future of judicial reform lies not in replacing paper with electronic documents but in redesigning judicial processes around structured, validated, and interoperable data. Such a transformation would significantly improve registry efficiency, reduce repetitive data entry, enhance judicial analytics, facilitate policy planning, and prepare the Indian judiciary for the next generation of AI-enabled legal technologies. The paper concludes with a conceptual framework and policy recommendations for implementing a phased transition from document-centric electronic filing to a comprehensive data-centric judicial information ecosystem.

Keywords: e-Courts, Digital India, e-Filing, Judicial Digital Transformation, PDF, Structured Data, Artificial Intelligence, Judicial Information Systems, National Judicial Data Grid, Digital Governance.

1. Introduction

The administration of justice has historically depended upon written records. For centuries, judicial institutions across jurisdictions have relied upon paper files containing pleadings, evidence, affidavits, orders, and judgments. Consequently, when courts worldwide embraced information technology, the earliest objective was naturally to replace paper with electronic documents. The transition from physical files to digital documents represented an important milestone in judicial modernization, promising improvements in accessibility, storage, and administrative efficiency. India has been no exception. Over the past two decades, the Indian judiciary has undertaken one of the world's largest judicial digitization initiatives through the e-Courts Mission Mode Project¹. The introduction of electronic filing, digital case management, virtual courts, online cause lists, electronic payment of court fees, and the National Judicial Data Grid has substantially transformed judicial administration². These reforms became particularly significant during the COVID-19 pandemic, when courts rapidly adopted virtual hearings and electronic filing mechanisms to ensure continuity of justice delivery³. Despite these commendable achievements, the existing architecture of electronic filing continues to mirror the traditional paper-based workflow. Litigants and advocates prepare pleadings in conventional document formats, convert them into Portable Document Format (PDF) files, digitally sign them, and upload them onto court portals. Registry officials thereafter examine these documents, manually extract relevant information, populate judicial databases, identify procedural defects, and process the matter for listing before the appropriate court. In essence, electronic filing has largely digitized the movement of paper without fundamentally transforming the underlying information architecture of judicial administration. This distinction between *digitization* and *digitalization* is both subtle and consequential. Digitization refers to the conversion of analogue information into digital form. Digitalization, by contrast, involves redesigning institutional processes around digital technologies to create

¹ Department of Justice, Government of India, *e-Courts Mission Mode Project*, available at: (last visited Aug. 25, 2026).

² National Informatics Centre, *e-Courts Services*, Government of India, available at: (last visited Aug. 25, 2026).

³ Supreme Court of India, *Guidelines for Court Functioning through Video Conferencing during COVID-19 Pandemic*, available at: (last visited Aug. 25, 2026).

new forms of efficiency, automation, and public value.⁴ While the Indian judiciary has successfully digitized judicial documents, it has not yet fully digitalized judicial information. The continued reliance upon PDFs as the principal repository of judicial information creates several structural limitations. First, PDFs are designed primarily for human reading rather than machine interpretation⁵. Although modern optical character recognition and natural language processing technologies can extract information from documents, such extraction remains computationally intensive, error-prone, and fundamentally unnecessary when structured information could have been captured at the point of filing. Secondly, registry officials continue to spend considerable time transferring information already available in pleadings into structured databases. Case categories, statutory provisions invoked, names of parties, addresses, advocates, reliefs claimed, limitation details, and jurisdictional particulars frequently require manual verification and entry. The judicial system therefore performs the same information processing exercise twice: once by the advocate while drafting the pleading and again by the registry while processing the uploaded document⁶. Thirdly, document-centric filing substantially limits the analytical capabilities of judicial databases. Contemporary case management systems primarily identify cases rather than understand them. They know that a petition exists but often cannot accurately determine the legal issues involved, the precise statutory provisions invoked, the relief sought, or the procedural stage reached without analysing the underlying documents. Consequently, judicial administrators cannot easily perform sophisticated policy analyses or generate evidence-based insights into litigation trends. The limitations become even more pronounced in the context of emerging artificial intelligence technologies. AI systems perform significantly better when information is available as structured, validated, and interoperable data⁷. By contrast, AI applications built upon millions of heterogeneous PDF documents must first reconstruct information that could have been captured directly during electronic filing. The resulting inefficiencies reduce accuracy, increase computational costs, and constrain the development of intelligent judicial assistance systems⁸.

⁴ OECD, *Going Digital: Shaping Policies, Improving Lives* (OECD Publishing 2019); OECD, *Digital Government Review of Brazil* (OECD Publishing 2018), distinguishing digitization from digital transformation in public administration.

⁵ Adobe, *Portable Document Format (PDF) Reference*, explaining the PDF format as a document-preservation and presentation format, available at: (last visited Aug. 25, 2026).

⁶ Department of Justice, Government of India, *e-Courts Mission Mode Project*, describing electronic filing, case information management and registry-related processing within the Indian judicial system, available at: (last visited Aug. 25, 2026).

⁷ OECD, *OECD Recommendation on Artificial Intelligence*, emphasising the importance of data quality and appropriate data governance for trustworthy AI systems, available at: (last visited Aug. 25, 2026).

⁸ National e-Governance Division, Ministry of Electronics and Information Technology, *National Data Governance*, Government of India, emphasising standardisation, interoperability, data quality and the use of data

This paper argues that the next phase of judicial reform must therefore move beyond electronic document management towards digital information management. The central proposition advanced herein is that judicial data not judicial documents should constitute the primary digital asset of the court system. Pleadings, affidavits, annexures, and evidence would continue to exist as authenticated documents; however, they should supplement rather than define the judicial information architecture. Such a transformation would align judicial administration with broader principles of digital governance. Across sectors including banking, taxation, healthcare, customs administration, and corporate regulation, governments increasingly collect structured transactional data directly from users instead of relying exclusively upon uploaded documents. Information entered once is subsequently reused across administrative processes, thereby reducing duplication, improving data quality, enabling automation, and facilitating evidence-based policymaking⁹. The judiciary, as one of the largest public information systems in India, stands to benefit substantially from adopting similar principles while respecting procedural fairness, judicial independence, and data protection requirements¹⁰. Accordingly, this paper examines the conceptual distinction between document-centric and data-centric judicial information systems. It analyses the limitations of the existing PDF-centric electronic filing architecture, explores comparative developments in digital governance, and proposes a conceptual framework for redesigning electronic filing around structured legal data. Rather than advocating the elimination of judicial documents, the paper proposes repositioning documents as evidentiary artifacts within a broader ecosystem in which structured information serves as the foundation for judicial administration, legal analytics, interoperability, and future AI-enabled court services.

2. Literature Review: From Electronic Documents to Data-Centric Justice

The digital transformation of judicial institutions has attracted increasing scholarly attention over the past two decades. The emergence of electronic filing systems, online dispute resolution, virtual hearings, digital evidence management, and artificial intelligence has

for Artificial Intelligence and evidence-based policymaking, available at: (last visited Aug. 25, 2026).

⁹ National e-Governance Division, Ministry of Electronics and Information Technology, *Meta Data and Data Standards*, Government of India, noting that standardised data elements facilitate the reliable reading, sorting, indexing, retrieval, communication and exchange of information between government systems, available at: (last visited Aug. 25, 2026).

¹⁰ The Digital Personal Data Protection Act, 2023, No. 22 of 2023, §§ 4–10, 11–17, India Code, Ministry of Law and Justice, Government of India, establishing statutory principles concerning lawful processing, obligations of data fiduciaries, rights of data principals and safeguards relating to digital personal data, available at: (last visited Aug. 25, 2026).

fundamentally altered the administration of justice in many jurisdictions. Existing literature has largely focused on improving access to justice through technology, enhancing judicial efficiency, and leveraging artificial intelligence for legal research and decision support. Comparatively less attention, however, has been devoted to the underlying architecture of judicial information systems and the distinction between document-centric and data-centric models of electronic filing.

2.1 Digitization and Digitalization in Public Administration

The distinction between *digitization* and *digitalization* has become a foundational concept in digital governance literature. Brennen and Kreiss define digitization as the technical process of converting analogue information into digital form, whereas digitalization refers to the broader transformation of organisational processes through digital technologies¹¹. Similarly, the Organisation for Economic Co-operation and Development (OECD) characterises digital government as a shift from simply automating existing procedures to redesigning institutions around digital data, interoperability, and citizen-centric services¹². This distinction is evident across sectors of public administration. Modern taxation systems require taxpayers to submit structured financial information through electronic forms, while supporting documents are uploaded only where necessary. Corporate regulatory systems similarly capture information in standardised digital formats that facilitate validation, analysis, and inter-agency exchange¹³. The principal administrative asset is therefore structured data rather than electronic documents. Judicial administration has generally adopted digital technologies more cautiously because of the unique requirements of procedural fairness, evidentiary integrity, and judicial independence. Nevertheless, the conceptual distinction between digitization and digitalization remains equally applicable to courts. Merely replacing physical files with electronic documents does not necessarily transform judicial processes if the underlying workflow continues to depend upon manual interpretation of documents.

¹¹ J. Scott Brennen & Daniel Kreiss, "Digitalization and Digitization," in *The International Encyclopedia of Communication Theory and Philosophy* (Klaus Bruhn Jensen, Robert T. Craig, Jefferson D. Pooley & Eric W. Rothenbuhler eds., Wiley Blackwell 2016).

¹² Organisation for Economic Co-operation and Development, *Digital Government Review of Sweden: Towards a Data-Driven Public Sector* 25–31 (OECD Publishing 2019), discussing the transition from traditional e-government towards digital government based on data, interoperability and redesigned public services.

¹³ World Bank Group, *GovTech Maturity Index 2022 Update: Trends in Public Sector Digital Transformation* 16–20 (World Bank 2022), discussing digital public infrastructure, interoperable government systems, shared data and the use of standardised digital information for public administration.

2.2 Evolution of Judicial Digital Transformation

The first generation of court computerisation primarily focused on administrative automation. Courts adopted case information systems, electronic cause lists, digital registers, and computerised case numbering to improve record management¹⁴. The second generation introduced electronic filing, online payment systems, digital signatures, and virtual hearings, thereby reducing dependence on physical interaction with court registries. The third and emerging generation seeks to incorporate artificial intelligence, predictive analytics, intelligent scheduling, automated document classification, and advanced judicial analytics. These technologies require structured, reliable, and interoperable data to function effectively. Several scholars have observed that technological advancement within courts has often replicated existing paper-based procedures rather than redesigning them. Richard Susskind argues that courts should not merely automate existing practices but should fundamentally rethink the manner in which legal services are delivered in a digital society¹⁵. His analysis emphasises that technology should transform institutional processes instead of simply reproducing analogue workflows in electronic form. Similarly, Katsh and Rabinovich-Einy argue that digital justice requires redesigning dispute resolution processes around information rather than documents, allowing technology to enhance efficiency without compromising procedural fairness¹⁶. Although their work focuses primarily on online dispute resolution, the underlying principle—that information architecture determines institutional capability—has broader implications for judicial administration.

2.3 The Indian Experience

India's judicial digital transformation has largely been driven through the e-Courts Mission Mode Project, initiated under the National e-Governance Plan¹⁷. The project has successfully computerised courts, established judicial service centres, implemented case information systems, enabled electronic filing, introduced virtual courts, and created the National Judicial Data Grid (NJDG). The Supreme Court's eCommittee has consistently

¹⁴ United Nations Office on Drugs and Crime, *E-Justice: Using Information and Communication Technologies in the Justice Sector* 7–12 (United Nations 2011), discussing the introduction of information and communication technologies into court administration, case management, electronic records and judicial information systems.

¹⁵ Richard Susskind, *Online Courts and the Future of Justice* 1–15 (Oxford University Press 2019).

¹⁶ Ethan Katsh & Orna Rabinovich-Einy, *Digital Justice: Technology and the Internet of Disputes* 3–18 (Oxford University Press 2017).

¹⁷ Department of Justice, Government of India, *National e-Governance Plan: e-Courts Mission Mode Project*, describing the objectives, phases and implementation of the e-Courts project for computerisation and modernisation of the Indian judiciary, available at: (last visited Aug. 25, 2026).

recognised technology as an instrument for improving access to justice¹⁸. Successive editions of the e-Courts Vision Document have emphasised interoperability, standardisation, digital records, and citizen-centric judicial services¹⁹. These initiatives represent significant achievements in judicial administration and have substantially increased transparency. However, the operational architecture of electronic filing remains predominantly document-oriented. Litigants continue to prepare pleadings in conventional formats, convert them into PDFs, digitally sign them, and upload them for registry scrutiny. Structured information relating to parties, statutory provisions, jurisdiction, limitation, relief sought, and procedural classification is either entered separately in limited fields or subsequently extracted by registry personnel. Existing policy documents discuss digitization extensively but devote comparatively limited attention to redesigning electronic filing around structured legal data as the primary repository of judicial information. Consequently, judicial databases frequently function as indexes pointing towards documents rather than comprehensive repositories of legal information²⁰.

2.4 Legal Informatics and Structured Legal Data

Legal informatics has long recognized the importance of structured legal information. Early developments concentrated on digital legal databases and computer-assisted legal research, while more recent scholarship has focused upon semantic legal technologies, legal knowledge representation, and machine-readable legislation. Researchers working in legal information retrieval have demonstrated that structured metadata significantly improves search accuracy, interoperability, and automated reasoning²¹. Contemporary artificial intelligence systems similarly rely upon structured datasets for training, validation, and explainability²². Within judicial administration, structured information enables automated jurisdictional checks,

¹⁸ Supreme Court of India, eCommittee, *Vision for Phase III of the e-Courts Project*, discussing technology-enabled courts, access to justice, digital information and transformation of judicial processes, available at: (last visited Aug. 25, 2026).

¹⁹ eCommittee, Supreme Court of India, *Draft Vision & Action Plan for Phase III of the e-Courts Project*, emphasizing an integrated digital judicial infrastructure, interoperability, standardisation and citizen-centric delivery of judicial services, available at: (last visited Aug. 25, 2026).

²⁰ National Judicial Data Grid, *About NJDG*, describing the NJDG as a database providing information relating to cases, judgments and judicial proceedings across Indian courts, available at: (last visited Aug. 25, 2026).

²¹ Kevin D. Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age* (Cambridge University Press 2017),

²² National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, 18–25 (2023), discussing data quality, representativeness, documentation, testing, validation and transparency as components of trustworthy AI systems, available at: (last visited Aug. 25, 2026).

intelligent scheduling, statistical analysis, workload distribution, and policy evaluation²³. By contrast, document-centric systems require repeated extraction of information that already exists within pleadings. The distinction may be illustrated by considering a civil suit. A PDF petition contains numerous identifiable data elements: names of parties, addresses, statutory provisions, limitation details, valuation, cause of action, and relief sought. A document-centric system stores these elements implicitly within the uploaded file, whereas a data-centric system captures each element explicitly as structured information while retaining the petition as evidentiary support. The latter approach allows validation at the time of filing, reduces registry workload, enhances data quality, and facilitates future analytical applications²⁴.

2.5 Artificial Intelligence and Judicial Data

The rapid development of artificial intelligence has intensified scholarly interest in judicial information systems. AI applications including intelligent document review, legal research assistants, automated classification, predictive analytics, and workflow automation increasingly depend upon high-quality structured data²⁵. Although advances in natural language processing have improved machine understanding of legal documents, extracting structured information from heterogeneous PDFs remains inherently less reliable than collecting validated information directly from users during electronic filing. AI therefore benefits most when documents supplement structured databases rather than substitute for them²⁶. The growing emphasis upon trustworthy AI also reinforces the importance of structured judicial information. Transparent, explainable, and accountable AI systems require traceable data sources and consistent data standards²⁷. A fragmented repository of uploaded documents presents greater challenges for verification, reproducibility, and algorithmic accountability

²³ European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*, paras. 11–14 (Council of Europe 2018), recognising the potential of AI and structured judicial data to improve the efficiency and quality of judicial systems while requiring appropriate safeguards, available at: (last visited Aug. 25, 2026).

²⁴ United Nations, *E-Government Survey 2022: The Future of Digital Government*, 31–38 (United Nations Department of Economic and Social Affairs 2022), discussing integrated digital government, interoperable data, once-only information provision, data quality and the reuse of information across administrative processes.

²⁵ Kevin D. Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age* 1–20 (Cambridge University Press 2017), discussing the use of computational techniques, machine learning and structured legal information in legal research and analysis.

²⁶ Harry Surden, “Artificial Intelligence and Law: An Overview,” 35 *Georgia State University Law Review* 1305, 1324–33 (2019), explaining the dependence of legal AI applications upon the availability, organisation and quality of underlying legal data.

²⁷ European Commission, *Ethics Guidelines for Trustworthy AI* 12–17 (High-Level Expert Group on Artificial Intelligence 2019), identifying transparency, traceability, accountability and appropriate data governance as essential requirements for trustworthy AI, available at: (last visited Aug. 25, 2026).

than an integrated data-centric information architecture.

2.6 Research Gap

The existing literature overwhelmingly recognises the importance of judicial digitisation, electronic filing, artificial intelligence, and digital governance. However, it generally treats electronic documents as the natural foundation of digital courts. Limited scholarly attention has been devoted to questioning whether documents themselves should remain the primary digital object within judicial information systems. This paper seeks to address that gap by advancing a conceptual distinction between **document-centric judicial architecture** and **data-centric judicial architecture**. It argues that the next stage of judicial reform should not merely improve electronic filing interfaces or automate document processing but fundamentally redesign judicial information systems around structured legal data. Such a shift would represent the transition from digitization to genuine digitalization within the Indian judiciary²⁸.

3. The Existing Architecture of Electronic Filing in the Indian Judiciary: The PDF Bottleneck

The electronic filing system introduced under the e-Courts Mission Mode Project marks a significant milestone in the modernization of the Indian judiciary²⁹. It has reduced dependence on physical records, facilitated remote filing by advocates and litigants, enabled digital scrutiny by court registries, and contributed to greater transparency in judicial administration. Nevertheless, the present architecture remains fundamentally document-centric. Electronic filing is primarily designed around the uploading, storage, and processing of Portable Document Format (PDF) files, while structured legal data occupies only a secondary role³⁰. This section argues that the principal bottleneck in India's electronic filing system is not technological but architectural. The existing workflow treats the uploaded PDF

²⁸ J. Scott Brennen & Daniel Kreiss, "Digitalization and Digitization," in *The International Encyclopedia of Communication Theory and Philosophy* (Klaus Bruhn Jensen, Robert T. Craig, Jefferson D. Pooley & Eric W. Rothenbuhler eds., Wiley Blackwell 2016), distinguishing the conversion of analogue information into digital form from the broader organisational transformation produced through digital technologies.

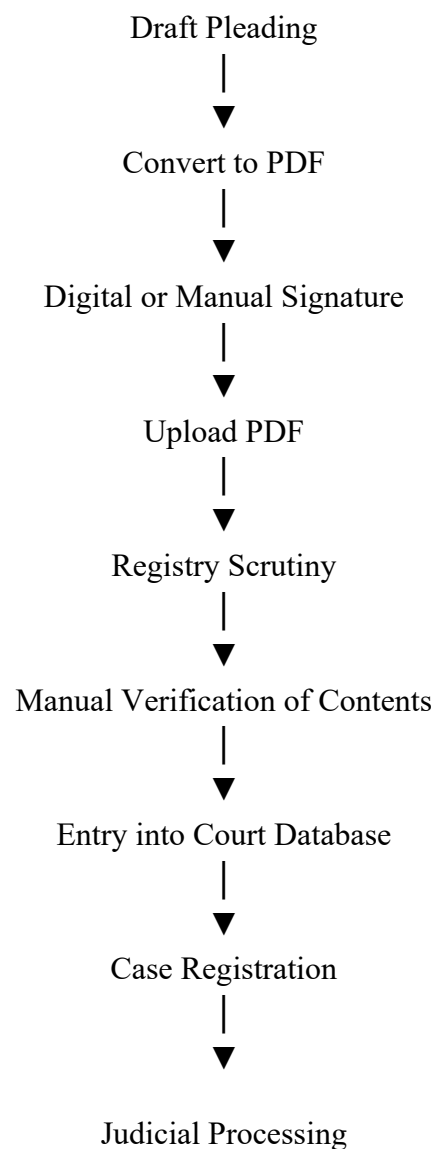
²⁹ eCommittee, Supreme Court of India, *e-Filing Services*, describing the electronic filing mechanism developed for filing cases and documents before Indian courts and its role in reducing physical interaction with court registries, available at: (last visited Aug. 25, 2026).

³⁰ eCommittee, Supreme Court of India, *e-Filing User Manual*, describing the electronic filing process, including preparation and uploading of pleadings and documents in prescribed electronic formats, available at: (last visited Aug. 25, 2026).

as the primary source of judicial information. Consequently, much of the information required for judicial administration must either be manually extracted from documents or repeatedly entered at different stages of case processing. The result is duplication of effort, increased administrative costs, inconsistent data quality, and limited opportunities for automation³¹.

3.1 The Existing e-Filing Workflow

A simplified representation of the prevailing electronic filing process is as follows:



³¹ World Bank, *GovTech Maturity Index 2022 Update: Trends in Public Sector Digital Transformation* 24–30 (World Bank 2022), discussing the administrative costs associated with fragmented information systems and the importance of integrated, interoperable and reusable government data for automation and improved public-service delivery.

Although electronic transmission replaces the physical movement of files, the logical workflow remains substantially unchanged from the traditional paper-based system. Registry officials continue to examine uploaded pleadings to verify compliance with procedural requirements, identify defects, determine case categories, confirm jurisdiction, record party particulars, and classify the matter before registration. In effect, the electronic filing portal functions primarily as a digital receiving counter rather than as an intelligent information system.

3.2 The PDF as the Primary Information Object

A PDF is designed to preserve the visual appearance of a document across different computing platforms. It ensures consistency in formatting, pagination, signatures, annexures, and evidentiary presentation. These characteristics make PDFs highly suitable for judicial records, where authenticity and fidelity to the original document are essential. However, a format optimised for human reading is not necessarily optimal for information processing.³² A petition may contain dozens of legally significant data elements, including:

- i. names of parties;
- ii. addresses and contact details;
- iii. advocate particulars;
- iv. statutory provisions invoked;
- v. jurisdictional facts;
- vi. limitation details;
- vii. valuation of the suit or claim;
- viii. reliefs sought;
- ix. details of previous litigation;
- x. particulars of annexures and supporting documents.

³² International Organization for Standardization, ISO 32000-2:2020, Document Management—Portable Document Format (PDF) 2.0, recognizing PDF as a document preservation and presentation standard rather than a structured database format.

Within the present system, most of these elements remain embedded within the uploaded document. Although certain metadata fields are captured separately during filing, they generally serve administrative indexing rather than comprehensive information management. Consequently, the registry frequently extracts information from documents that already exist in electronic form. This process is analogous to receiving an electronic income tax return as a scanned PDF and subsequently requiring tax officials to manually populate the revenue database from that document. Such an approach would be regarded as inefficient in virtually every other domain of digital governance.

3.3 Duplicate Data Entry: The Hidden Administrative Burden

One of the least examined consequences of the PDF-centric architecture is the duplication of data entry. During preparation of pleadings, advocates already identify parties, statutory provisions, addresses, causes of action, reliefs, valuation, limitation, and jurisdiction. Once converted into PDFs, however, this structured information effectively becomes locked inside an electronic document. Registry personnel subsequently repeat much of the same exercise by reading the uploaded document and entering corresponding details into the Case Information System (CIS) or related databases. The consequence is a duplication of administrative effort. Information that has already been created digitally is re-created through manual intervention because the filing architecture captures documents rather than data. This duplication has several adverse consequences:

- i. increased processing time during scrutiny;
- ii. greater likelihood of transcription errors;
- iii. inconsistent data across judicial databases;
- iv. additional workload for registry personnel;
- v. delays in defect rectification; and
- vi. reduced scope for automated validation.

Viewed collectively, these inefficiencies constitute what this paper describes as the **PDF Bottleneck**.

3.4 The PDF Bottleneck: An Architectural Problem

The expression "**PDF Bottleneck**" is used in this paper to describe the situation in which judicial information becomes operationally inaccessible because it is stored primarily within uploaded documents rather than as structured digital data³³. The bottleneck does not arise because PDFs are technologically deficient. Indeed, PDFs remain indispensable for preserving pleadings, affidavits, evidence, judgments, and signed legal documents³⁴. Rather, the problem arises when PDFs are expected to perform functions for which they were never designed. A judicial information system performs numerous tasks beyond document preservation. It must classify cases, allocate judicial workload, identify related proceedings, monitor timelines, generate statistics, facilitate appellate review, support legal research, and increasingly enable intelligent digital services. These functions require structured information that can be validated, queried, analysed, and exchanged electronically³⁵. When the PDF itself becomes the principal information repository, every subsequent digital process depends upon either manual interpretation or computational extraction of embedded information. The architecture therefore places an unnecessary informational barrier between data creation and data utilisation.

3.5. Comparison with Other Digital Public Systems

The limitations of document-centric architecture become clearer when compared with other public digital platforms in India. The Goods and Services Tax Network (GSTN) captures invoice details, tax liabilities, and transaction information directly as structured digital data. Supporting documents are required only in exceptional circumstances.³⁶ Likewise, the Income Tax Department's electronic filing system records financial information through structured forms while allowing documentary uploads only where necessary.³⁷ Corporate filings under the Ministry of Corporate Affairs similarly rely upon structured electronic forms that facilitate

³³ Tim Berners-Lee, James Hendler & Ora Lassila, "The Semantic Web," 284 *Scientific American* 34, 35–43 (2001).

³⁴ International Organization for Standardization, *ISO 32000-2:2020, Document Management—Portable Document Format—Part 2: PDF 2.0*, specifying the technical framework for the Portable Document Format as a means of representing and preserving electronic documents.

³⁵ World Wide Web Consortium (W3C), *Data on the Web Best Practices*, explaining the importance of machine-readable, interoperable, accessible and reusable data for data processing and exchange, available at: (last visited Aug. 25, 2026).

³⁶ Goods and Services Tax Network (GSTN), *GST Returns and Electronic Filing Framework*, Government of India.

³⁷ Income Tax Department, Government of India, *e-Filing Portal User Manuals and Return Filing Utilities*.

automated validation and regulatory oversight.³⁸ These systems illustrate an important design principle: documents support data; they do not replace data. The judiciary, however, continues to rely predominantly upon uploaded pleadings as the principal source of operational information. Registry officials, therefore, undertake information processing that has already been completed by advocates during document preparation. This contrast demonstrates that judicial digitization has progressed more slowly than digital transformation in other sectors of public administration—not because courts are resistant to technology, but because judicial information architecture continues to be organised around electronic documents rather than structured legal data.

3.6 Consequences for Judicial Administration

The implications of the PDF-centric model extend beyond registry operations. First, judicial databases become dependent upon manually entered metadata, limiting both completeness and reliability. Secondly, policy-makers cannot easily perform sophisticated empirical analyses regarding litigation patterns because many legally significant variables remain embedded within unstructured documents³⁹. Thirdly, interoperability with external government databases becomes more difficult because structured data standards are absent or inconsistently applied⁴⁰. Fourthly, artificial intelligence applications must expend significant computational resources extracting information that could have been collected directly during filing⁴¹. Finally, future technological innovations become constrained by historical design choices. Instead of building intelligent judicial services upon structured information, developers are compelled to reconstruct that information retrospectively from millions of heterogeneous PDF documents. These limitations do not diminish the remarkable achievements of the e-Courts Project. Rather, they indicate that the next stage of reform should focus less upon digitising additional documents and more upon redesigning judicial information architecture itself. The transition from digitization to digitalization therefore

³⁸ Ministry of Corporate Affairs, Government of India, *MCA21 e-Governance Project: Electronic Forms and Digital Filing Framework*.

³⁹ Richard A. Posner, *How Judges Think* 105–15 (Harvard University Press 2008), discussing the importance of empirical information concerning judicial decision-making and institutional performance for understanding the functioning of courts.

⁴⁰ European Commission, *European Interoperability Framework—Implementation Strategy*, COM(2017) 134 final, emphasising interoperability, common data standards and the ability of public-sector information systems to exchange and reuse information, available at: (last visited Aug. 25, 2026).

⁴¹ OECD, *OECD Framework for the Classification of AI Systems*, 15–22 (OECD Publishing 2022), discussing the role of data inputs, data quality and data processing in the performance and deployment of AI systems, available at: (last visited Aug. 25, 2026).

requires recognising that while documents remain indispensable as legal evidence, they should no longer constitute the primary operational database of the judiciary⁴².

4. From Documents to Data: A Data-Centric Architecture for Judicial Information Systems

The preceding section demonstrated that the principal limitation of the existing e-Filing system lies not in the use of PDF documents per se but in treating those documents as the primary repository of judicial information. This section advances the central thesis of this paper: the future of judicial digital transformation depends upon recognising that documents are evidence, whereas data is infrastructure. This distinction represents more than a technological refinement. It requires a fundamental shift in the conceptual architecture of judicial administration. Courts will continue to preserve pleadings, affidavits, annexures, and judgments as authenticated legal documents. However, the operational functioning of the judicial system should increasingly depend upon structured legal data captured, validated, and managed throughout the life cycle of a case.

4.1 The Principle: Documents are Evidence; Data is Infrastructure

Every judicial proceeding generates two distinct but interconnected assets. The first is the legal document. Pleadings, written statements, affidavits, documentary evidence, and judgments constitute the official legal record. They preserve the exact language employed by parties and the court, possess evidentiary value, and remain indispensable for adjudication and appellate review⁴³. The second is the judicial information contained within those documents. Every petition identifies parties, advocates, statutory provisions, jurisdictional facts, limitation periods, reliefs claimed, procedural stages, and numerous other data elements that enable the administration of justice. Traditional paper systems necessarily embedded these two assets within the same physical file. Digital systems are not subject to that constraint. They can preserve the original document while simultaneously maintaining a structured representation

⁴² Council of Europe, European Commission for the Efficiency of Justice (CEPEJ), *European Judicial Systems—CEPEJ Evaluation Report 2024*, discussing the continuing digital transformation of judicial systems, information systems and the use of technology to improve judicial efficiency and accessibility, available at: (last visited Aug. 25, 2026).

⁴³ National Archives of India, *Digital Records Management Guidelines*, Ministry of Culture, Government of India, discussing the authenticity, integrity, reliability and preservation of digital records as official records, available at: (last visited Aug. 25, 2026).

of its legally significant information⁴⁴. The failure to distinguish between these two functions has resulted in electronic filing systems that preserve documents effectively but manage information inefficiently. This paper therefore proposes the following architectural principle:

“Legal documents should remain the authoritative evidentiary record, whereas structured judicial data should become the primary operational resource of the digital court.”

Such a distinction neither diminishes the legal significance of pleadings nor alters established procedural rules. Rather, it recognises that digital governance requires information systems capable of processing structured data while preserving authenticated documentary records.

4.2 Data Capture at the Point of Filing

A data-centric filing system should collect structured information directly from litigants or their authorised representatives at the time of filing. Instead of merely uploading a completed petition, the filing process would require users to enter standardised information through validated electronic forms. The uploaded pleading would continue to accompany the filing but would primarily serve evidentiary and archival purposes. Illustratively, a civil complaint may generate structured fields such as:

- i. names and identities of parties;
- ii. addresses and jurisdictional particulars;
- iii. statutory provisions invoked;
- iv. valuation of the suit;
- v. limitation calculations;
- vi. nature of relief sought;
- vii. previous or connected litigation;
- viii. subject matter classification;

⁴⁴ International Council on Archives, *Principles and Functional Requirements for Records in Electronic Office Environments*, emphasising the preservation of authentic electronic records together with the metadata and contextual information necessary for their management and future use, available at: (last visited Aug. 25, 2026).

- ix. details of annexures;
- x. advocate particulars.

Many of these details are already entered in some form within existing systems. The proposed model differs in treating these fields not as ancillary metadata but as the authoritative operational dataset from which judicial processes are conducted. Validation can occur immediately at the point of filing. Mandatory fields cannot be omitted. Jurisdictional inconsistencies can be flagged automatically. Duplicate party names can be identified. Deficiencies may be corrected before the registry undertakes scrutiny. Consequently, registry personnel shift from repetitive data transcription towards legal verification and quality assurance.

4.3 The Lifecycle of Judicial Data

A judicial case is not a static document but a continuously evolving process. As proceedings progress, additional structured information is generated:

- i. service of summons;
- ii. filing of written statements;
- iii. framing of issues;
- iv. interlocutory applications;
- v. evidence stages;
- vi. witness examinations;
- vii. interim orders;
- viii. final judgments;
- ix. appeals;
- x. execution proceedings.

Each procedural event generates valuable information capable of supporting judicial administration. Under a document-centric model, these developments are primarily recorded through uploaded orders and documents. Under a data-centric model, each event updates the

structured case record while simultaneously preserving the underlying judicial documents. The result is a living judicial database capable of reflecting the procedural history of every case with far greater precision than conventional document repositories.

4.4. Interoperability and the Digital Justice Ecosystem

Modern public administration increasingly depends upon interoperable digital platforms capable of securely exchanging verified information⁴⁵. A data-centric judicial architecture would facilitate standardised interaction with other governmental systems where legally permissible and institutionally appropriate. Illustrative examples include:

- i. verification of company information through the Ministry of Corporate Affairs;
- ii. validation of property identifiers from digital land records;
- iii. exchange of information with prisons and correctional institutions;
- iv. electronic transmission of warrants and judicial orders;
- v. integration with forensic laboratories;
- vi. authenticated communication with police agencies;
- vii. digital payment systems for court fees and compensation⁴⁶.

Such interoperability should be governed by statutory safeguards, judicial oversight, data minimisation principles, and robust cybersecurity measures⁴⁷. The objective is not unrestricted information sharing but controlled, accountable, and legally authorised exchange of structured judicial information. A document-centric architecture makes such interoperability considerably

⁴⁵ United Nations Department of Economic and Social Affairs, *United Nations E-Government Survey 2024: Accelerating Digital Transformation for Sustainable Development*, discussing integrated digital public infrastructure, interoperability, secure data exchange and digitally enabled public services (United Nations 2024), available at: (last visited Aug. 25, 2026).

⁴⁶ Department of Justice, Government of India, *e-Courts Services*, describing digital judicial services including e-payment of court fees, fines and other judicial transactions, available at: (last visited Aug. 25, 2026).

⁴⁷ Digital Personal Data Protection Act, 2023, No. 22 of 2023, 4–10, India Code, Ministry of Law and Justice, Government of India, providing the statutory framework governing processing of digital personal data and obligations relating to its protection, available at: (last visited Aug. 25, 2026).

more difficult because external systems cannot readily interpret unstructured pleadings without extensive manual intervention or computational processing⁴⁸.

4.5 Enabling Artificial Intelligence Responsibly

The emergence of artificial intelligence has intensified interest in judicial technologies. AI applications may assist courts in document classification, legal research, scheduling, translation, transcription, workload management, and statistical analysis⁴⁹. However, the effectiveness of these technologies depends fundamentally upon data quality. Artificial intelligence performs best when trained and deployed upon structured, validated, and consistently organised information⁵⁰. Although advances in natural language processing have improved the ability of machines to interpret legal documents, extracting reliable information from millions of heterogeneous PDFs remains both computationally expensive and susceptible to error. A data-centric judicial architecture significantly reduces these limitations. Rather than reconstructing legal information from documents after filing, AI systems can rely upon information that has already been validated through judicial processes. Importantly, this paper does not advocate replacing judicial reasoning with algorithmic decision-making⁵¹. Adjudication necessarily involves interpretation, discretion, constitutional values, and human judgment that cannot be reduced to computational rules. Instead, structured judicial data should support administrative intelligence, not judicial substitution. Artificial intelligence should assist registries, court administrators, researchers, and judges by improving access to information while preserving the decisional independence of the judiciary⁵².

⁴⁸ World Wide Web Consortium (W3C), *Data on the Web Best Practices*, discussing machine-readable data, interoperability, data reuse, metadata and the technical requirements for effective exchange and processing of data across systems, available at: (last visited Aug. 25, 2026).

⁴⁹ European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*, paras. 31–38 (Council of Europe 2018), identifying potential applications of AI in judicial systems, including processing judicial decisions, legal research, case management and access to justice, available at: (last visited Aug. 25, 2026).

⁵⁰ National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, 18–27 (2023), emphasising data quality, representativeness, reliability and appropriate data governance as important considerations for trustworthy AI systems, available at: (last visited Aug. 25, 2026).

⁵¹ Council of Europe, *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*, Principle Five, emphasising that AI applications in judicial systems must respect the principle of user control and must not replace judicial decision-making or individual judicial discretion, available at: (last visited Aug. 25, 2026).

⁵² United Nations Educational, Scientific and Cultural Organization (UNESCO), *Recommendation on the Ethics of Artificial Intelligence*, paras. 24–29 (2021), emphasising human oversight, accountability, transparency and the protection of human agency in the deployment of AI systems, available at: (last visited Aug. 25, 2026).

4.6 A Conceptual Comparison

The distinction between the present architecture and the proposed model may be summarised as follows:

Present PDF-Centric Model	Proposed Data-Centric Model
Document is the primary information object.	Structured data is the primary operational object.
Registry extracts information from pleadings.	Information is captured once during filing.
Metadata is limited.	Comprehensive structured judicial dataset.
Manual scrutiny dominates routine verification.	Automated validation supports registry scrutiny.
Analytics depend upon document extraction.	Analytics operate directly on validated data.
AI must interpret heterogeneous documents.	AI utilises structured and traceable information.
Documents drive administration.	Documents preserve evidence; data drives administration.

The proposed model should not be understood as replacing documents with databases. Rather, it recognises that judicial administration requires both. Documents preserve legal authenticity; structured data enables digital governance.

4.7 Towards the Next Generation of Digital Courts

The history of technological reform demonstrates that genuine innovation rarely results from merely automating existing practices. Instead, transformation occurs when institutions reconsider the assumptions underlying established processes⁵³. The first generation of judicial reform computerised paper records. The second generation enabled electronic transmission of documents. The third generation now emerging must redesign judicial information architecture around structured, interoperable, and reusable legal data. India's judiciary has already established the institutional foundations necessary for such a transformation through the e-Courts Mission Mode Project, digital infrastructure, and the National Judicial Data Grid. The next phase is therefore not one of incremental digitisation but of architectural digitalization⁵⁴.

⁵³ Richard Susskind, *Tomorrow's Lawyers: An Introduction to Your Future* 1–12 (2d ed., Oxford University Press 2017), arguing that technological change in legal services requires fundamental reconsideration of established methods rather than merely automating conventional legal processes.

⁵⁴ Organisation for Economic Co-operation and Development, *The OECD Digital Government Policy Framework: Six Dimensions of a Digital Government*, OECD Public Governance Policy Papers No. 02, 5–12 (2020), distinguishing digitally enabled government transformation from the mere adoption of information

The transition from document-centric to data-centric judicial administration is neither a technological luxury nor an abstract academic proposition. It represents the logical evolution of electronic justice systems capable of supporting transparency, efficiency, interoperability, empirical policy-making, and the responsible adoption of artificial intelligence⁵⁵.

5. Policy Recommendations and a Roadmap for Transitioning to a Data-Centric Judicial Information System

The transition from a document-centric to a data-centric judicial information system cannot be achieved merely by introducing new software or expanding storage capacity. It requires a comprehensive redesign of judicial processes, procedural rules, technical standards, and institutional practices. Importantly, the proposed transformation should be evolutionary rather than disruptive. Courts must continue to preserve the evidentiary integrity of judicial documents while progressively adopting structured data as the principal operational resource. This section proposes a phased implementation strategy capable of achieving that objective.

5.1 Guiding Principles for Reform

Any redesign of judicial information architecture should be guided by five foundational principles. First, digital systems must preserve procedural fairness. Technology should simplify judicial administration without diminishing the procedural safeguards available to litigants⁵⁶. Secondly, the authenticity and evidentiary value of judicial documents must remain unaffected. Structured data should complement, not replace, the legally binding record⁵⁷. Thirdly, information should be captured once and reused throughout the lifecycle of a case. Repeated manual transcription of identical information should become the exception rather

technology and emphasising government-wide data-driven and user-centred transformation, available at: (last visited Aug. 25, 2026).

⁵⁵ Council of Europe, European Commission for the Efficiency of Justice (CEPEJ), *European Judicial Systems—CEPEJ Evaluation Report 2024*, discussing digitalisation of judicial systems, information and communication technologies, data-driven judicial administration and emerging technological applications in justice systems, available at: (last visited Aug. 25, 2026).

⁵⁶ Council of Europe, European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*, Principle I, requiring respect for fundamental rights, including the right to a fair trial, in the use of AI and digital technologies within judicial systems, available at: (last visited Aug. 25, 2026).

⁵⁷ United Nations Commission on International Trade Law (UNCITRAL), *Model Law on Electronic Commerce with Guide to Enactment*, arts. 8–10 (1996), recognising the legal significance of electronic records and establishing principles concerning their evidentiary value, authenticity and integrity, available at: (last visited Aug. 25, 2026).

than the norm⁵⁸. Fourthly, interoperability should be achieved through open standards rather than proprietary technological solutions. Judicial independence requires that court information systems remain capable of long-term maintenance and technological evolution without dependence upon particular vendors⁵⁹. Finally, every technological intervention must comply with constitutional principles relating to equality, access to justice, transparency, privacy, and accountability⁶⁰. These principles ensure that digital transformation strengthens rather than compromises the administration of justice.

5.2 Phase I: Standardisation of Judicial Data

The first phase should focus upon identifying the minimum structured information required for every category of judicial proceeding. The Supreme Court's eCommittee, in consultation with High Courts, judicial academies, the Bar, and technical experts, may develop a National Judicial Data Standard (NJDS) prescribing uniform data elements across Indian courts. Illustratively, every case could include standardised fields relating to:

- i. identity of parties;
- ii. advocates;
- iii. statutory provisions invoked;
- iv. case category;
- v. relief sought;
- vi. valuation;
- vii. territorial and pecuniary jurisdiction;
- viii. limitation particulars;

⁵⁸ Organisation for Economic Co-operation and Development, *The OECD Digital Government Policy Framework: Six Dimensions of a Digital Government*, OECD Public Governance Policy Papers No. 02, 8–12 (2020), identifying data-driven public-sector transformation and the reuse of information as core characteristics of mature digital government, available at: (last visited Aug. 25, 2026).

⁵⁹ European Commission, *European Interoperability Framework—Implementation Strategy*, COM(2017) 134 final, emphasising open specifications, interoperability, sustainability and avoidance of technological lock-in within public-sector information systems, available at: (last visited Aug. 25, 2026).

⁶⁰ *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1, recognising privacy as a constitutionally protected fundamental right under Article 21 of the Constitution of India and emphasising the constitutional requirements of legality, necessity and proportionality in relation to measures affecting privacy.

- ix. connected proceedings;
- x. stage of litigation;
- xi. disposal outcome.

Uniform standards would significantly improve consistency across jurisdictions while facilitating interoperability among judicial databases. Importantly, such standardisation concerns administrative information and does not interfere with judicial discretion regarding pleadings, evidence, or adjudication.

5.3 Phase II: Structured Electronic Filing

Once common data standards are established, electronic filing portals should progressively transition from document-upload platforms to structured filing systems. Instead of beginning with a PDF upload, litigants would first complete validated electronic forms containing legally significant information. The pleading itself would then be uploaded as an authenticated supporting document. This sequence reverses the present architecture.

Currently:

Document → Registry → Data

Proposed model:

Validated Data → Supporting Documents → Registry Verification

The registry would therefore verify legal compliance instead of repeatedly extracting routine information already supplied electronically. Such an approach is consistent with the design of modern digital public platforms in taxation, customs administration, corporate regulation, and electronic procurement.

5.4 Phase III: Intelligent Registry Management

The registry performs one of the most information-intensive functions within judicial administration. A structured data architecture would permit the gradual introduction of intelligent administrative assistance, including:

- i. automatic identification of incomplete filings;
- ii. validation of mandatory procedural requirements;
- iii. detection of duplicate or related proceedings;
- iv. verification of jurisdictional particulars;
- v. limitation alerts;
- vi. automated scheduling recommendations;
- vii. statistical reporting;
- viii. electronic workflow management.

These functions would assist registry personnel while leaving all adjudicatory decisions entirely under judicial control. The objective is administrative efficiency rather than judicial automation.

5.5 Phase IV: Interoperable Digital Justice

The long-term objective should be the development of an interoperable judicial information ecosystem. Such an ecosystem would permit secure and legally authorised exchange of structured information among judicial institutions and other public authorities. Potential areas include:

- i. prison administration;
- ii. forensic laboratories;
- iii. legal services authorities;
- iv. mediation centres;
- v. police agencies;
- vi. land registration authorities;
- vii. corporate registries;
- viii. insolvency institutions;
- ix. tribunal systems.

Interoperability should occur only within clearly defined statutory frameworks incorporating authentication, encryption, audit trails, access controls, and independent oversight. Digital integration must never compromise judicial confidentiality or individual privacy.

5.6 Institutional and Legal Challenges

The proposed transition inevitably presents significant institutional challenges. The first concerns change management. Judicial officers, advocates, registry personnel, and litigants have developed extensive experience with document-based procedures. Successful implementation therefore requires sustained capacity building and professional training⁶¹. Secondly, technical standardisation across diverse jurisdictions presents considerable complexity. Indian courts vary substantially in caseloads, procedural practices, linguistic diversity, and technological infrastructure. Uniform standards must therefore retain sufficient flexibility to accommodate legitimate institutional differences⁶². Thirdly, privacy and data governance require careful consideration. A data-centric architecture necessarily involves increased collection of structured information. Appropriate legal safeguards concerning access, retention, correction, anonymisation, and security should therefore accompany technological reforms, consistent with applicable data protection legislation and constitutional jurisprudence⁶³. Finally, reforms should proceed incrementally through pilot projects, continuous evaluation, and consultation with all stakeholders⁶⁴. Digital transformation within the judiciary succeeds when technological innovation is accompanied by institutional confidence.

5.7 The Constitutional Dimension

The proposed reforms should not be understood merely as administrative

⁶¹ United Nations Department of Economic and Social Affairs, *United Nations E-Government Survey 2024: Accelerating Digital Transformation for Sustainable Development*, emphasising institutional capacity, digital skills, workforce preparedness and organisational change as prerequisites for successful digital transformation, available at: (last visited Aug. 25, 2026).

⁶² Organisation for Economic Co-operation and Development, *Digital Government Review of India: Accelerating the Digital Transformation of the Public Sector*, discussing the need for common digital standards and interoperable systems while accounting for institutional and administrative diversity (OECD Publishing 2020).

⁶³ *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1; Digital Personal Data Protection Act, 2023, No. 22 of 2023, 4–17, India Code, Ministry of Law and Justice, Government of India, establishing constitutional and statutory principles relevant to the lawful processing and protection of personal data, available at: (last visited Aug. 25, 2026).

⁶⁴ World Bank, *GovTech Maturity Index 2022 Update: Trends in Public Sector Digital Transformation*, emphasising iterative digital transformation, institutional capacity, user engagement, monitoring and continuous improvement in public-sector technology projects (World Bank 2022).

improvements. Efficient judicial information systems directly advance constitutional values. Article 14 of the Constitution requires fairness and non-arbitrariness in public administration⁶⁵. Accurate and consistent judicial information contributes to equal treatment of similarly situated litigants. Article 21, as interpreted by the Supreme Court, encompasses access to fair, timely, and effective justice⁶⁶. Reducing unnecessary administrative delays through better information management strengthens rather than weakens procedural fairness. Similarly, open judicial data appropriately anonymised where necessary enhances transparency, empirical legal research, and informed public policy⁶⁷. The transition towards data-centric judicial administration therefore supports broader constitutional objectives relating to the rule of law, institutional accountability, and access to justice⁶⁸.

5.8 Reimagining the Judicial Record

Perhaps the most significant implication of this paper is conceptual rather than technological. Historically, the judicial record has been understood as a collection of authenticated documents. Digital courts invite a broader understanding. The judicial record should instead be conceived as comprising two inseparable but distinct components:

1. **the authoritative documentary record**, preserving the exact pleadings, evidence, and judicial orders; and
2. **the structured judicial information record**, enabling administration, analytics, interoperability, and future digital services.

Neither component can substitute for the other. Documents establish legal rights and obligations. Structured data enables the judicial institution to administer those rights efficiently. Recognising this distinction marks the transition from electronic record

⁶⁵ *E.P. Royappa v. State of Tamil Nadu*, (1974) 4 SCC 3, 38–39, holding that arbitrariness is antithetical to equality and that Article 14 strikes at arbitrary State action.

⁶⁶ *Hussainara Khatoon (I) v. Home Secretary, State of Bihar*, (1980) 1 SCC 81, 89–90, recognising the right to speedy trial as an essential component of the right to life and personal liberty under Article 21 of the Constitution of India.

⁶⁷ *Swapnil Tripathi v. Supreme Court of India*, (2018) 10 SCC 639, recognising the importance of greater accessibility and transparency in judicial proceedings and holding that proceedings of constitutional importance may, subject to appropriate safeguards, be made accessible through live-streaming.

⁶⁸ *Anita Kushwaha v. Pushap Sudan*, (2016) 8 SCC 509, recognising access to justice as a fundamental constitutional entitlement and emphasising that access includes an effective adjudicatory mechanism capable of providing timely and meaningful justice.

management to digital judicial governance.

VI. Conclusion: Towards a Data-First Doctrine of Judicial Digitalization

The digital transformation of the Indian judiciary has reached an important point of inflection. Over the past two decades, substantial investments have been made in court computerisation, electronic filing, virtual hearings, digital case management, and online judicial services⁶⁹. These initiatives have significantly improved accessibility, transparency, and administrative efficiency. They have also demonstrated the judiciary's willingness to embrace technology while preserving the integrity of judicial processes. Yet, as this paper has argued, the present architecture of electronic filing remains rooted in the logic of paper-based administration. Although documents now travel electronically rather than physically, the underlying workflow continues to revolve around the preparation, submission, scrutiny, and storage of PDF documents. Information essential for judicial administration frequently remains embedded within those documents, requiring repeated extraction and manual processing. This paper has described this phenomenon as the **PDF Bottleneck**. The bottleneck is not created by the Portable Document Format itself, which remains an indispensable medium for preserving pleadings, evidence, affidavits, and judicial orders. Rather, the bottleneck arises because judicial information systems continue to treat the document as the primary operational object instead of treating structured legal data as the principal resource for judicial administration. The distinction between **digitization** and **digitalization** therefore assumes critical importance. Digitization replaces paper with electronic documents. Digitalization redesigns institutional processes around structured information, interoperability, automation, and evidence-based governance. While the Indian judiciary has achieved considerable success in digitizing judicial records, the next phase of reform requires the digitalization of judicial information itself.

The central contribution of this paper has been to propose a conceptual distinction between **document-centric judicial architecture** and **data-centric judicial architecture**. A document-centric architecture performs judicial administration through uploaded documents. Data is extracted after filing, often manually, to populate administrative databases. By contrast, a data-centric architecture captures validated legal information at the point of filing, enabling judicial administration to operate directly upon structured information while preserving

⁶⁹ Department of Justice, Government of India, *Annual Report 2023–24*, discussing the implementation and development of the e-Courts Mission Mode Project, computerisation of courts, e-filing, virtual courts, digital judicial services and related technological infrastructure, available at: (last visited Aug. 25, 2026).

documents as the authoritative evidentiary record. The implications of such a transition extend beyond registry efficiency. Structured judicial data enables improved case management, more reliable judicial statistics, empirical legal research, interoperable digital governance, responsible deployment of artificial intelligence, and evidence-based judicial policy⁷⁰. It also reduces duplication of effort, improves data quality, and creates opportunities for intelligent administrative assistance without compromising judicial independence or adjudicatory discretion. Accordingly, this paper advances what may be described as the **Data-First Doctrine of Judicial Digitalization**. The doctrine rests upon four normative propositions. **First**, the primary digital asset of a court should be structured judicial information rather than uploaded documents. **Secondly**, authenticated judicial documents should continue to constitute the legally authoritative evidentiary record and remain indispensable to adjudication. **Thirdly**, information required for judicial administration should be captured once, validated at the point of origin, and reused throughout the lifecycle of the proceeding, thereby eliminating unnecessary duplication and enhancing administrative efficiency⁷¹. **Fourthly**, digital technologies, including artificial intelligence, should augment judicial administration by supporting registries, research, scheduling, and analytics, while leaving adjudicatory functions exclusively under human judicial control⁷². Collectively, these principles provide a conceptual framework for the next generation of judicial digital transformation.

The proposed doctrine should not be understood as advocating the elimination of PDF documents or conventional pleadings. Courts necessarily depend upon authentic documentary records. Rather, it argues that digital courts should distinguish between **evidence** and **information**. Documents preserve legal authenticity; structured data enables institutional intelligence. Both are essential, but they perform different functions within a modern judicial information system. The significance of this distinction extends beyond India. Many jurisdictions that have introduced electronic filing continue to organise digital justice around electronic documents. Consequently, the conceptual framework developed in this paper may

⁷⁰ Kevin D. Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age* 1–20 (Cambridge University Press 2017), discussing the importance of computationally processable legal information for legal analytics, information retrieval and technology-assisted legal work.

⁷¹ Organisation for Economic Co-operation and Development, *The OECD Digital Government Policy Framework: Six Dimensions of a Digital Government*, OECD Public Governance Policy Papers No. 02, 8–12 (2020), identifying data-driven government, interoperability and the reuse of information as fundamental characteristics of mature digital government, available at: (last visited Aug. 25, 2026).

⁷² Council of Europe, European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*, Principle Five, emphasising that AI applications in judicial systems must remain under human control and must not impair judicial independence or replace judicial decision-making, available at: (last visited Aug. 25, 2026).

assist comparative discussions concerning judicial information architecture, legal informatics, and digital governance in other common law systems. For the Indian judiciary, the transition from a PDF-centric system to a data-centric judicial information ecosystem represents more than a technological upgrade. It is a shift in institutional philosophy. The objective is not merely to make paperless courts, but to create **data-enabled courts** capable of delivering justice more efficiently, transparently, and intelligently while remaining firmly grounded in constitutional values, procedural fairness, and judicial independence⁷³. The future of digital justice will therefore not be determined by the number of documents uploaded to judicial servers. It will be determined by how effectively courts transform those documents into structured knowledge that supports the administration of justice. The next chapter in the evolution of the Indian judiciary is thus not simply the digitization of judicial records; it is the **digitalization of judicial intelligence**.

⁷³ *Anita Kushwaha v. Pushap Sudan*, (2016) 8 SCC 509; *Swapnil Tripathi v. Supreme Court of India*, (2018) 10 SCC 639, recognising access to justice and transparency as important dimensions of the constitutional functioning of the judicial system.