HEURISTICA
DISCOVERY COUNSEL LLP

When Evidence Goes Digital: Reliability Risks in Arbitration

December 1, 2026

At the Forefront of Discovery

When Evidence Goes Digital: Reliability Risks in Arbitration

Introduction

The integrity of evidence is fundamental to the legitimacy of any adjudicative process. Digital evidence and generative AI are changing how arbitrators and counsel confront the challenges of ensuring that evidence is disclosed and weighed fairly. Records are now created, stored, and transmitted not only in digital form, but increasingly with generative and agentic AI. With this transformation comes new questions about reliability, authenticity, integrity, and evidentiary weight.[1]

These developments present particular challenges for arbitration. Arbitration tribunals have the authority to determine the admissibility, relevance, materiality and weight of evidence, and to establish procedures for the disclosure and examination of evidence. That discretion is subject to fundamental principles of procedural fairness, including the opportunity for each party to present its case and respond to the evidence against it. Concerns about reliability and authenticity of admitted evidence can affect confidence in the arbitral process and its outcomes, and in some cases could lead to an award being set aside.

The methods used to evaluate evidence must keep pace with the realities of the digital environment. Recognizing and adapting to the risks associated with digital and generative AI evidence can preserve arbitration credibility and enforceability of awards in a changing world. Business records are no longer confined to paper or documents that can be converted to paper. They are now drawn from complex systems, often as exports or data extracts, and increasingly shaped by algorithms, AI-driven processes, and agentic AI. Routine communications, such as email and messaging platforms, frequently point to internal or external links, embedded content, or dynamic sources that are not part of the record itself. Expert reports, once based solely on disclosed documents and individual professional judgement, may now incorporate automated analyses or outputs from generative AI tools.

The evolution of evidence away from paper documents (upon which our law, procedures and practices have been based) complicate how reliability and completeness can and should be assessed. In many common law countries, through legislation or case law, computerized business records are presumed to be reliable and admissible[2].

The presumption of reliability and, further, the tendency to accept it as proof of system integrity or substantive accuracy, is based on an outdated assumption that computer systems and software are bug free. This assumption is dangerous. Some of the dangers of assuming reliability are well illustrated in the

[1] Briefly: Authenticity asks, "Is this document what the tendering party claims it is?"; Integrity asks, "Has the electronic record or system maintained the relevant information without material alteration?"; Reliability asks, "Can the information in the record be trusted as an accurate representation of the events or transactions?"; and evidentiary weight asks, "How much probative weight should the tribunal place in that information?"

[2] See for example sections 31.1 to 31.8 of the *Canada Evidence Act*. In England, there is a common law principle that computer evidence should be presumed reliable.

largest widespread miscarriage of justice in United Kingdom legal history, commonly referred to as the Post Office Horizon Scandal, which caused hundreds of individuals to be wrongfully prosecuted. With the growing use of generative AI and agentic AI, which are known to hallucinate and provide unexpected output, the risk of relying on inaccurate or unreliable evidence is only increasing.

Arbitration does not need rigid evidentiary rules to respond to this risk. The flexibility of arbitration allows arbitrators to tailor disclosure obligations to the procedural fairness requirements of each matter, including challenges posed by any digital or AI-generated evidence. This paper examines these issues and considers how arbitrators and counsel can balance efficiency and proportionality with the need for reliable evidence and fair outcomes.

Presumption of Reliability

The presumption of reliability of computer evidence is intended to promote efficiency, particularly in cases involving massive volumes of electronic records. It avoids turning every digital document into a mini trial on system integrity and admissibility. However, the presumption has proven to have significant negative consequences.

Proof of system integrity does not necessarily establish the substantive accuracy of the information recorded by that system. A system may reliably preserve inaccurate information, and a properly authenticated document may still contain errors. This distinction is particularly important where business records are generated by complex software systems whose operations, limitations and underlying data are not readily accessible to the opposing party. It becomes even more complex with generative and agentic AI. Parties who control the systems that generate the evidence relied upon have a strong evidentiary advantage, while opposing parties are left to challenge reliability without necessarily having the access or expertise to do so.

Where there are reasonable challenges to reliability, those who rely on digital evidence should disclose enough metadata, audit logs, and system documentation to allow meaningful testing of reliability. In exceptional circumstances, where less intrusive testing cannot adequately address a material concern, access to underlying systems and source code may be justified, particularly where disputed digital system records form the backbone of liability or damage assessments.

Lessons from the UK Horizon Scandal

The United Kingdom's Post Office Horizon Scandal illustrates what can happen when faulty digital evidence is presumed to be reliable. A link to the Post Office Horizon IT Inquiry is **here**.

For more than 15 years, the Fujitsu Horizon computer system generated financial shortfalls in local Post Office branches that were treated as evidence of theft and fraud. Hundreds of postmasters and others were prosecuted and convicted. Many were financially ruined, some were imprisoned, and there were several suicides.

The landmark appellate decision, *Hamilton & Others v Post Office Ltd*[3], was the first to overturn convictions. The Court of Appeal found that serious defects in the Horizon system, together with the Post Office's failures to investigate and disclose known reliability concerns, had rendered the prosecutions an abuse of process.

When a later inquiry revealed extensive flaws in the software, it became clear that the problem was not only technical but evidentiary. It was determined that the software had systemic software bugs, despite its owner actively maintaining it was infallible, calling into question all the convictions.

The Horizon Inquiry highlights the problems that can be created when a party without access to the computer system has the burden to prove the opposing party's software/computer system/SaaS platform was faulty, which is an impossible task without access to system-level information and an expert. This problem is compounded by proprietary systems and the introduction of various forms of AI and agentic AI into business systems.

The reliability issues that arose in the Post Office prosecutions can arise in any proceeding. Parties routinely rely on extracts from enterprise systems such as financial ledgers, project reports and, increasingly, AI transcripts of meetings and agentic AI. When the opposing party challenges the integrity of the evidence with indicia of unreliability, system and document integrity should no longer be presumed. Indicia of problems can include missing metadata, inconsistent timestamps, inadequate preservation or insufficient audit capability.

If reasonable questions about integrity and reliability are raised and remain unanswered, the probative value of the evidence is impacted. Depending on the circumstances, the possible evidentiary consequences include less weight being given to the evidence, an adverse inference being drawn or a refusal to admit the record into evidence. If the digital evidence is no longer available where there was an obligation to preserve (for example, in the context of a dispute or regulatory proceeding or through contractual obligation) adverse inferences are appropriate in many jurisdictions and mandatory in Canada.[4]

Arbitral procedures and confidence are preserved not by ignoring questions of integrity and reliability but by dealing with them in a reasoned and proportionate way.

CGSB's "Electronic Records as Documentary Evidence"

There are no binding principles for standards of system integrity. The Canadian General Standards Board (CGSB)'s "Electronic Records as Documentary Evidence"[5] provides a practical framework for assessing the integrity and reliability of electronic records, helping to mitigate the risks of relying on presumptions of system integrity.[6] This standard provides a vocabulary and framework to test reliability that is principled

[3] [2021] EWCA Crim 577
[4] See *SS&C Technologies Canada Corp. ("SS&C")* v. *Bank of New York Mellon Corp. ("BNY")*, 2026 SCC 29
[5] Originally titled CGSB-72.34-2017, the standard was updated in 2024 and is now CAN/CGSB-72.34-2024.
[6] There are other standards for digital records that may be used as reference for considering indicia of unreliability. See British Standard BS 10008, which can be sourced here: BS 10008-1:2020 – TC.

and practical. While it is not binding, compliance with the standard reinforces an argument that the electronic records in question are reliable, accurate, authentic, and therefore capable of being admissible as legal evidence. Arbitrators may consider compliance with the standard as evidence of proper records management.[7]

Key features of the standard include:

a. Auditability: records should be supported by metadata and audit trails that demonstrate their origin, integrity, and chain of custody.[8]

b. System Controls: organizations should have policies and safeguards in place to prevent unauthorized alteration.[9]

c. Retention and Preservation: records must remain accessible and trustworthy over time, including during litigation holds.[10]

Auditability requires that records can be traced through metadata and audit logs to confirm their creation, modification, access, and export. This allows both the tribunal and the opposing party to test whether the record is consistent with the narrative advanced. System controls address whether and how records could have been altered. The tribunal does not need to review enterprise policy in the abstract, but it should understand, for the records at issue, who had permission to alter or delete them and whether such actions would have been recorded. Preservation ensures that once a dispute was anticipated, records were held intact.

For counsel challenging digital evidence, grounding requests in a standard like the CGSB standard or BS 10008 can demonstrate that they are not speculative demands but principled attempts to test reliability. Where warranted, ordering production of metadata and audit trails in line with a standard transforms presumed reliability into something verifiable. It gives arbitrators and counsel a benchmark against which to measure whether records are trustworthy, rather than relying on presumptions.

Attached, as Appendix A, are two charts setting out some points to consider in a challenge to an IT system (system integrity) or in a challenge to individual records (records integrity).

Expert Reports and AI

Expert reports increasingly incorporate the use of artificial intelligence and generative systems. Opaque systems may generate outputs that appear authoritative but cannot be tested. How can it be determined if the AI inappropriately influenced the judgment of the expert?

[7] See CEA, s. 31.5
[8] CGSB-72.34-2024, Policy 6
[9] *ibid*., Policy 4
[10] *ibid*., Policy 9

When experts rely on AI, the risk is not novelty but the inability to test the method. An expert who feeds selective excerpts into a generative system and reports its output as analysis is offering conclusions that cannot be cross-examined upon. Without particulars of scope, inputs, method, and validation, the tribunal cannot know whether the opinion is actually the opinion of the expert or whether it reflects little more than a machine's unverified output. The effect of AI on the judgment of experts was recently studied by Anthropic. On October 27, 2026, Anthropic produced a study titled "*Who's Aligned to Whom? How AI Agents Shape Experts' Research Judgment Calls*" that found using a fictional advisory program that intentionally leaned the user to particular decisions, a professional chose options that matched the AI preference 86% of the time. Even the experts were not aware of how their opinions were being influenced by the AI.

If the inner workings of an AI system are explainable, then it is easier to get people to trust the results and find the system to be reliable. However, AI systems often, and increasingly, generate output through opaque processes that are not easily explained, even by experts. This is known as the "black box" problem.

That said, meaningful challenge does not always require technical explainability. Full explainability is often impossible, prohibitively costly, and not very helpful to adjudicators. Instead, arbitrators should assess AI evidence the way that the courts have treated other accepted 'black boxes', such as pharmaceuticals whose mechanisms of action are unknown or forensic expertise based on experience. The focus should be on inputs and outputs, not on inner workings. Was the methodology peer-reviewed? What is the error rate? Has the output been independently verified? In law, what matters is not explanation but justification, that is, whether the result can be grounded in legal norms and tested under the rules of evidence.

Where AI-generated evidence is materially relied upon, tribunals should consider whether sufficient information has been provided concerning the inputs, methodology, limitations and validation of the output to permit meaningful scrutiny. The appropriate safeguards will depend on the nature of the evidence and the purpose for which it is tendered.

While there is not yet a Canadian case on point, a recent American appellate case was faced with whether evidence from an AI-assisted investigative tool was properly excluded by the lower court. In *State v. Carr*[11], the Ohio Court of Appeals considered the disclosure obligations associated with evidence generated by CyberCheck, a proprietary AI-assisted investigative tool. The trial court had excluded the evidence after the prosecution failed to produce the underlying data, algorithms and software necessary for the defence to independently evaluate its reliability. On appeal, the exclusion order was reversed, not because the appellate court accepted the reliability of the technology, but because the trial court had failed to consider less severe measures to address the disclosure deficiency. These included compelling production from the third-party technology provider, permitting independent examination of the software, or conducting a hearing to determine whether the methodology satisfied applicable evidentiary standards. The decision illustrates both the difficulties associated with testing proprietary AI-generated evidence and the importance of proportionate procedural responses.

[11] 2024-Ohio-4471

Challenges to reliability should not automatically require disclosure of source code or complete technical explainability. The appropriate inquiry is what information or testing is reasonably necessary to evaluate the particular conclusions being advanced. Tribunals should consider targeted disclosure, independent testing and other procedural safeguards before resorting to exclusion of the evidence. Procedural fairness does not require prying open the black box or excluding the evidence; it requires ensuring that the outcomes of these systems can be scrutinized against legal standards. That being said, exclusion of evidence may still be appropriate where meaningful testing is impossible or evidence is only partially available.

Increasingly, experts rely on AI tools to prepare portions of their analysis. The reliability of an expert opinion depends on a sufficiently reliable and verifiable factual foundation. If AI is used, arbitrators must be assured that the expert's conclusions are rooted in actual evidence. The maxim "garbage in, garbage out" is apt here. If the AI system is trained on or fed flawed, incomplete, or inadmissible data, then the resulting analysis will be equally compromised. Ensuring that expert reports remain anchored in admissible, verifiable evidence is therefore critical to maintaining their probative value.

Consider an example where an opposing party's expert report quotes excerpts from alleged productions, but the entire documents from which the excerpts allegedly came are not specified. First, we are faced with the question of whether the excerpts actually do come from documents in the production set, and if so, which ones. Further, even if they do come from the productions, we query whether the AI tool used by the expert had the proper inputs: Did the expert merely feed the algorithm selected quotes from the documents instead of the entirety of each admissible document? What other documents were used? If the inputs can be challenged, the output may be challenged.

Attached as Appendix B is a table which sets out some of the issues that you might want to consider when AI-generated evidence is proffered and challenged.

Witness Preparation and Evidence

There is an emergent recognition that generative AI can unintentionally impact witness evidence. Recent research from MIT media labs, "*AI-Implanted False Memories*", found the following:

> [A] consistent pattern: AI systems, whether through suggestive questioning, visual manipulation, or conversational misdirection, possess an unprecedented capacity to shape and distort human memory. This work establishes that the integration of AI into sensitive contexts such as legal proceedings, therapeutic settings, and everyday information consumption requires careful consideration of these memory manipulation risks.

Concern about a witness using AI to prepare to provide evidence in a criminal proceeding in the UK resulted in the English Court of Appeal[12] finding that a witness using AI to prepare could give rise to an abuse of process finding:

[12] R v FGD [2026] EWCA Crim 918

> All witnesses, whether for the prosecution or defence, should be firmly discouraged from resorting to the use of AI to prepare for giving evidence. Doing so could run the risk of unfairness up to the point, in an extreme case, where the court may have to consider excluding evidence.

In Canada, under most jurisdictions' civil law procedural rules, a witness who uses AI to prepare for pre-trial examinations under oath could conceivably be compelled to produce the AI materials relied upon to prepare for the examinations.

Preservation and Spoliation of Digital Evidence

In common law jurisdictions there is an obligation to preserve relevant evidence, and intentional spoliation of evidence carries risks of sanctions in various forms. In Canada, the importance of preserving digital evidence was recently reinforced by the Supreme Court of Canada in *SS&C Technologies Canada Corp. v. Bank of New York Mellon Corp.*[13] The case involved the non-production of electronic data relevant to determining the extent to which proprietary financial information had been improperly distributed. Despite receiving a preservation notice in anticipation of litigation, the defendant bank failed to preserve the relevant data, maintaining that the allegations against it were without merit.

The Supreme Court clarified the doctrine of spoliation in Canada, confirming that it arises where relevant evidence is intentionally destroyed, altered or concealed with a view to affecting litigation. The importance of this decision lies in the clear statement that intent to affect the litigation can be inferred from a failure to preserve.

Once spoliation is established, the court must presume that the missing evidence would have been unfavourable to the party responsible, although that presumption is rebuttable. Where it is not rebutted, the court must draw mandatory adverse inferences that meaningfully address the evidentiary gap created by the missing evidence. The nature and scope of those inferences depend on the circumstances and the prejudice caused.

Although not every failure to preserve evidence will amount to spoliation, the decision reinforces the importance of implementing effective preservation measures when a dispute is reasonably contemplated. For arbitration, the principles provide persuasive guidance when considering the consequences of missing digital evidence, including whether adverse inferences are warranted. More broadly, the decision underscores that a party should not benefit from its deliberate destruction of evidence that would otherwise permit meaningful scrutiny of its case.

Hyperlinked Attachments

Hyperlinked attachments, sometimes called "modern attachments", are files that are shared by inserting a link into an email or chat message to a file stored in the cloud (such as OneDrive, Google Drive, or SharePoint), rather than attaching a static copy of the file. Where you have a traditional attachment, that attachment is preserved as part of the email itself, creating a fixed snapshot of the document the moment

[13] 2026 SCC 29

it was sent. Conversely, a hyperlinked attachment points to a live document that can be changed, moved, or deleted at any time. What appears to be a minor change in technology can create major evidentiary problems if the data is not collected properly and promptly.

Unlike traditional attachments which are embedded and are automatically captured and indexed along with the individual email, hyperlinked documents require a specific approach to be collected. This means that relevant content may be overlooked if collection workflows do not properly identify and collect linked files. As a result, critical evidence may never be considered, despite being within the control of the producing party. This is compounded by collecting using keywords which would not search the content of linked attachments.

Even where the linked attachments are collected, problems can arise. A link points to a live file that may have been altered, moved, or deleted after the message was sent. This means the version collected later may differ from the one originally shared. This makes it far more difficult to prove what the recipient saw and when, and it also breaks the usual family relationship between a message and its attachments that is critical for preserving context.

There are potential solutions. For example, certain Microsoft licences may now allow for the ability to collect the hyperlinked document as it was initially shared or all versions of the linked document. However, these capabilities are subject to limitations. Recovering the contemporaneous version depends on factors such as the organization's licensing, retention policies, version history and whether the relevant functionality was configured when the document was shared. The key point is that it is crucial that the issue be explicitly addressed early by the parties and in disclosure agreements or procedural orders.

In practice, counsel should inquire early about their clients' communication technologies and how they are used to ensure that relevant hyperlinked documents are identified, preserved and collected. Where hyperlinked documents are not collected or their contemporaneous versions cannot be established, questions may arise concerning their integrity and reliability. The party relying on such evidence should be expected to address credible concerns. If the concerns remain unresolved, the evidentiary foundation for admissibility may be undermined, or the weight accorded to the record may be substantially diminished.

For arbitrators, a dispute about the production of hyperlinked attachments raises two distinct questions: what information must be preserved and collected, and how much additional effort is justified to reconstruct the relationship between a communication and the document it references.

While there is not yet a developed body of Canadian jurisprudence addressing disclosure obligations for hyperlinked attachments, American decisions provide useful guidance. In an early decision on the topic, *Nichols v. Noom Inc.* (2021)[14], the court declined to treat hyperlinked documents as conventional email attachments. By contrast, in *In re Uber Technologies, Inc. Passenger Sexual Assault Litigation* (2024)[15], the

[14] 2021 WL 948646 (S.D.N.Y. 2021). It must be noted that in 2021 the use of hyperlinked attachments was fairly novel and not as common.
[15] 2024 WL 1772832 (N.D. Cal. 2024)

court interpreted the parties' disclosure protocol to include modern attachments, while recognizing the technological difficulties associated with recovering contemporaneous versions.[16]

More recently, in *United Association National Pension Fund v. Carvana Co.* (2026)[17], the court ordered targeted forensic collection of contemporaneous versions of hyperlinked documents, while limiting the scope of that exercise in recognition of the substantial costs involved. Conversely, in *Yotta Technologies Inc. v. Evolve Bank & Trust* (2026)[18], the court declined to require production of linked documents merely because they were referenced by hyperlinks, emphasizing the need for a particularized justification.

In *Eli Lilly and Company v. Adonis Health, Inc.* (2026)[19], the court further clarified that an agreed protocol governing hyperlinked attachments does not relieve a party of its broader obligation to collect responsive records from relevant cloud-based systems. In that case, the parties had agreed that a file linked to a produced document does not have to be produced as part of the document family. However, the linked repositories clearly contained relevant information, and the party was ordered to collect and produce responsive, nonprivileged documents from each of five repositories.

As technology changes, it has also become much easier to preserve and capture internal links to SharePoint or other collaborative sources and, for some internal corporate systems, it can be as easy as preserving the emails. Conversely, links to external third-party sources are not typically collectable with the emails.

Disclosure obligation should depend on the circumstances, the evidentiary significance of the linked documents, the applicable agreement between the parties, and the proportionality of the proposed collection efforts. Further, it is important to address hyperlinked attachments expressly in disclosure agreements to avoid evidentiary gaps and disputes over collection obligations.

Conclusion

The disputes that arbitrators and parties now face are not decided on paper ledgers and signed letters, but on system-generated reports, AI-assisted analysis, and records whose integrity depends on metadata and audit trails rather than ink and signatures. Rules to short-circuit extensive testing of systems have been relied on without question, namely, the presumptions of system integrity and the business records exception. These principles were not designed with modern evidentiary realities in mind. Statutory presumptions of system integrity and the traditional rules governing business records can facilitate the admission of electronic evidence, but neither necessarily establishes the substantive accuracy of the information recorded. The Post Office Horizon Scandal is a cautionary warning: misplaced confidence in system integrity can distort justice for years before the error becomes visible.

---

[16] Subsequent orders in *Uber* (2025) addressed the production of linked documents and associated metadata. See https://www.govinfo.gov/app/details/USCOURTS-cand-3_23-md-03084/USCOURTS-cand-3_23-md-03084-26/context?utm

[17] 2026 WL 80189 (D. Ariz. 2026).

[18] 2026 WL 172502 (N.D. Cal. 2026)

[19] No. 25-cv-03536-JST (LB), Discovery Order (N.D. Cal. Oct. 16, 2026)

The challenge for arbitration is not to abandon efficiency or impose rigid evidentiary requirements. It is to recognize when the nature of the evidence or a credible challenge to its reliability warrants closer examination. Standards such as those from CGSB and BS offer vocabulary for integrity and reliability, but their application might raise the question of whether they enhance the legitimacy of the arbitral process or, instead, risk importing rigid formalism into a process valued for its flexibility.

These are not purely technical matters. They go to the heart of evidence integrity and arbitral legitimacy. If tribunals rely on digital records that have not been properly tested when reasonably challenged, they risk undermining confidence in the award and in arbitration as a forum. If counsel do not press for disclosure in principled ways, they risk conceding the very foundation of their case or making disproportionate requests. And if arbitrators do not address reliability concerns, efficiency may be purchased at the expense of justice.

The primary question is not about what the law requires, but about what arbitration needs to maintain its credibility in the digital age. When you next sit as arbitrator or act as counsel, consider the following:

- When the reliability of digital evidence is credibly challenged, how can the tribunal balance proportionality with the need to ensure that the awards rest on evidence that has been subjected to appropriate scrutiny?
- As AI-generated analysis and system-level evidence become routine, should a tribunal rely on presumptions and weight adjustments, or is it time to articulate firmer expectations of disclosure and integrity?

These are not abstractions. They are the issues that will continue to face tribunals and counsel and define whether arbitration continues to deliver decisions that parties, and the broader legal community, can trust.

Appendix A

Arguments to challenge system integrity and records integrity

A. System integrity

System Integrity Arguments (IT system as a whole)
Intermittent compliance: Standard requires continuous compliance with policies and procedures; partial or reactive compliance undermines system integrity.
Authentication failures: No reliable process to prove that the system consistently produces authentic records (e.g., no records management (RM) manual or IT system documentation).
Weak records management program: Lack of a formal RM policy/manual, unclear Records Officer authority, or inconsistent classification and retention procedures.
Inadequate audit trails: Inadequate audit trails may limit the ability to reconstruct material events or changes.
Security/control failures: Weak access controls, encryption lapses, or poor backup validation allow unauthorized access or tampering.
Unreliable conversion/migration: Poorly documented or unverifiable digital conversions/migrations create uncertainty that system processes preserve record accuracy.
Emerging tech risks: Use of AI, social media, or mobile devices without risk assessments creates uncertainty about trustworthiness.
Legal hold failures: Inability to suspend destruction/disposition when litigation is anticipated undermines confidence in overall system preservation.

B. Records integrity

Records Integrity Arguments (individual records)
Altered content: Evidence that the record has been modified, corrupted, or otherwise changed from its original state may defeat record integrity if the alteration was not legitimate, authorized or immaterial.
Missing or inconsistent metadata: Missing or inconsistent metadata may raise questions about a record's origin, identity or integrity.
Multiple versions without version control: Absence of documented versioning procedures raises doubt about which copy is authoritative.
Breaks in chain of custody: Gaps in the documented chain of custody may warrant further inquiry into preservation and handling.
Inconsistent business use: If an organization does not actually rely on the record in its ordinary course of business, its reliability as evidence is questionable.
Lack of contemporaneity: Records not created, received, or stored within a reasonable time after events may lack reliability.
Data entry errors: Records entered outside normal business processes, or without documented procedures, may be inaccurate.
Incomplete preservation: Incomplete preservation may make it more difficult to establish integrity or test the record's reliability.

Appendix B

Challenges to AI-generated evidence

Category	Issues for Challenge	Key Considerations
Threshold Admissibility	• Relevance • Authentication	– Does the AI output address a material issue? – Who sponsors and authenticates it?
Reliability of the AI System	• Explainability / transparency • Validation and testing • Bias and training data • Updates / versioning • Auditability	– Can the process be explained or is it a 'black box'? – Has the system been validated and error-rates measured? – Was the training data biased or incomplete? – Was the same version used consistently? – Is there a reproducible audit trail?
System & Records Integrity	• Metadata preservation • Tampering risks	– Is metadata intact to show creation and alterations? – Could the output have been manipulated or “deep faked”?
Collection & Discovery Concerns	• Reproducibility • Completeness • Methodology	– Can the result be replicated with the same inputs? – Were all relevant inputs disclosed? – Were collection and preprocessing transparent and defensible?
Fairness & Prejudice	• Right to challenge • Undue prejudice • Access to comparable tools	– Can the opposing party meaningfully cross-examine the AI? – Will the novelty/objectivity of AI overawe the trier of fact?[20] – Does one party have exclusive access to proprietary systems?
Legal & Ethical Standards	• Compliance with statutory frameworks	– Does the evidence align with Canada Evidence Act s.31.5 and provincial rules?

[20] AI systems often present results in a highly technical, statistical, or authoritative form. Because of this, there is a danger that decisionmakers may treat the output as inherently more reliable than it is. This is sometimes referred to as “automation bias”.

Category	Issues for Challenge	Key Considerations
	• Case law precedent • Professional responsibility	– How have courts treated AI/digital evidence to date? – Has counsel met duties of competence and candour?