HEURISTICA
DISCOVERY COUNSEL LLP

When Evidence Goes Digital: Reliability Risks in Arbitration

At the Forefront of Discovery

Introduction

Arbitrators and counsel have always confronted the challenge of ensuring that evidence is disclosed fully and weighed fairly. Those fundamental responsibilities remain unchanged. What is changing, however, is the nature of the evidence itself. Records are now created, stored, and transmitted in digital form, and with that shift come new questions about reliability, authenticity, and integrity. The traditional skills of arbitration professionals continue to be essential, but the processes and tools used to evaluate evidence must keep pace with the realities of the digital environment. Recognizing and adapting to these risks is not about abandoning established practice, but about preserving arbitration's credibility and effectiveness in a world where the character of evidence is evolving.

Business records are no longer confined to paper. They are now drawn from complex systems, often as exports or data extracts, and increasingly shaped by algorithms and even AI-driven processes. Expert reports, once grounded solely in disclosed documents and professional judgment, may now incorporate automated analyses or outputs from generative tools. Even routine communications such as email, messaging platforms, and collaboration tools frequently point to external links, embedded content, or dynamic sources that are not part of the record itself. Each of these shifts complicates how reliability and completeness can be assessed.

The tragedy of the Horizon prosecutions in the United Kingdom provides a cautionary example of what can happen when computer-generated business records are presumed to be reliable. In Canada, the *Canada Evidence Act* ("CEA") framework creates such a presumption. However, the Canadian General Standards Board (CGSB)'s "Electronic Records as Documentary Evidence"[1] can be a valuable tool to address the risk created by presuming reliability.

These issues can directly impact confidence in the arbitral process and decisions. For counsel, the issues can shape what disclosure to seek, how to frame evidentiary challenges, and how to present records in a form that an arbitral tribunal can accept and rely upon. This paper addresses those questions, not to propose rigid evidentiary codes, but to show that the flexibility of arbitration can tailor procedures and orders to safeguard fairness and evidentiary reliability.

[1] Originally titled CGSB-72.34-2017, the standard was updated in 2024 and is now CAN/CGSB-72.34-2024.

Lessons from the UK Horizon Scandal

The United Kingdom's Post Office Horizon scandal illustrates what can happen when digital evidence is presumed to be reliable. For more than 15 years, the Horizon accounting software generated discrepancies that were treated as proof of theft and fraud. Hundreds of sub-postmasters were wrongfully prosecuted and convicted. Many were financially ruined; some were imprisoned; and at least four died by suicide. When later inquiry revealed extensive flaws in the software, it became clear that the problem was not only technical but evidentiary. It was subsequently determined that the software produced errors calling into question all the convictions.

The leading appellate decision, *Hamilton & Others v Post Office Ltd*[2], quashed dozens of convictions, finding that the prosecutions were an abuse of process because the reliability of the Horizon system had never been properly investigated.

The Horizon scandal highlights the injustice of placing the burden on the individuals to prove the computer system was faulty, which is an impossible task without access to technical expertise and system-level information. This burden can be compounded by the use of proprietary document storage systems and the introduction of various forms of AI into business systems.

Arbitration is exposed to the same dynamic faced by the courts in the Horizon scandal. Parties routinely rely on extracts from enterprise systems such as financial ledgers, project reports, and communication logs. When the opposing party challenges the integrity of the evidence with specific concerns, for example, that metadata is missing, that timestamps are inconsistent, or that preservation measures were inadequate, integrity should not be presumed, it should be demonstrated. If reasonable integrity questions are raised and remain unanswered, the probative value of the evidence is diminished. The tribunal is entitled to adjust weight accordingly and, where refusal to disclose system-level information or metadata appears to be strategic, to draw adverse inferences. Arbitral procedures and confidence are preserved not by ignoring questions of integrity but by resolving them in a structured and proportionate way.

***CEA* and CGSB's "Electronic Records as Documentary Evidence"**

While the CEA is not binding in arbitration, the presumption it establishes is often implemented, even if not by direct reference. Sections 31.1 to 31.8 of the CEA establish a shortcut for the authentication of electronic evidence. Section 31.2(1) provides that a record generated or stored by an electronic system that "operated properly" at the relevant time is presumed to be authentic and accurate. Once invoked, this presumption shifts the burden to the other side to prove the record is unreliable. This presumption was one of the underlying failures identified in the Horizon scandal.
The presumption 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 authenticity. However,

[2] [2021] EWCA Crim 577

the presumption also has significant consequences. 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 some cases, this will include direct access to the system and software coding. This is especially important in cases where disputed digital system records form the backbone of liability or damage assessments.

The CGSB's "Electronic Records as Documentary Evidence" standard[3] provides a vocabulary and framework to test reliability that is principled and practical. While it is not binding, it can provide guidance for ensuring that electronic records are reliable, accurate, authentic, and therefore capable of being admissible as legal evidence. It emphasizes proving system integrity and record integrity to satisfy rules of authenticity, best evidence, and the business records exception in the CEA. Arbitrators may consider compliance with the standard as evidence of proper records management.[4]

Key features of the standard include:

a. Auditability: every record should be supported by metadata and audit trails that demonstrate its origin, integrity, and chain of custody.[5]

b. System Controls: organizations should have policies and safeguards in place to prevent unauthorized alteration.[6]

c. Retention and Preservation: records must remain accessible and trustworthy over time, including during litigation holds.[7]

[3] Originally titled CGSB-72.34-2017, the standard was updated in 2024 and is now CAN/CGSB-72.34-2024.
[4] See CEA, s. 31.5
[5] CGSB-72.34-2024, Policy 6
[6] *ibid*., Policy 4
[7] *ibid*., Policy 9

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 the CGSB standard 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 the 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, like the one in the CEA.

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.

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 grounded or whether it reflects little more than a machine's unverified output.

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.

Canadian criminal justice research reinforces the problematic nature of the AI black box. The Law Commission of Ontario[8] has warned that black box algorithms raise risks to fundamental Charter rights, including s. 7 (life, liberty, and security of the person) and s. 11(d) (fair trial guarantees), if accused persons cannot meaningfully challenge AI-generated evidence. That said, meaningful challenge does not always require technical explainability.

[8] Law Commission of Ontario, "AI at Trial and on Appeal", AI in Criminal Justice Project, Paper 4, April 2025 [See https://www.lco-cdo.org/en/our-current-projects/crimai/]

Full explainability is often impossible, prohibitively costly, and not very helpful to the arbitration panel. 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. For that reason, admissibility should require transparency about inputs and outputs, disclosure of system limitations, and evidence of testing and error rates. These are the safeguards that allow counsel to probe reliability in practice.

While there is not yet a Canadian case on point, a recent American appellate case recognized the limits of explainability. In *State v Carr*[9], the Court of Appeals in the Ninth Judicial District found that the AI-generated evidence the state tried to proffer was not admissible because it failed to meet evidentiary standards. The decision recognized the challenge of AI's black box but did not focus on it. Instead, the court found that while full technical explainability may be unrealistic, due process requires that AI evidence be capable of meaningful challenge through disclosure of methodology, inputs, and limitations. Without such safeguards, admitting AI evidence risks violating the accused's rights.

In arbitration, similar risks apply. Parties may insist on disclosure of metadata, audit logs, and system documentation so that AI-generated or system-based evidence can be tested. Procedural fairness does not require prying open the black box; it requires ensuring that the outcomes of these systems can be scrutinized against legal standards.

In dispute resolution, we see challenges with AI-generated evidence incorporated into expert reports. Increasingly, experts rely on AI tools to prepare portions of their analysis. The reliability of an expert opinion depends on the foundation of admissible evidence. If AI is used, arbitrators must be assured that the expert's conclusions are rooted in actual evidence rather than unverifiable algorithmic outputs. 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.

[9] 2024-Ohio-4471

In one case example, 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.

There are rule changes underway in Ontario and US Federal court where it is anticipated that expert evidence rules will be amended to address authenticity and reliability when AI is used.

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.

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 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 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, however, solutions that parties can implement to address this issue. For example, certain Microsoft licences now allow for the ability to collect the hyperlinked document as it was initially shared or all versions of the linked document. There are other options to properly collect hyperlinked documents. The key point is that it is crucial that the issue be explicitly addressed early by the parties and in the disclosure agreements.

While there is no Canadian case law regarding hyperlinked attachments yet, there are some instructive decisions from the US. In an early decision from 2021, *Nichols et al v. Noom Inc. et al*[10] , the court held that hyperlinks in emails were not automatically attachments under the agreed protocol, prioritizing proportionality over completeness. By contrast, more recently, in *In re Uber Techs., Inc. Passenger Sexual Assault Litig.* [Uber][11], the court interpreted "attachments" broadly to include modern attachments, though it acknowledged technological limits on producing contemporaneous versions. Similarly, in the *StubHub Refund Litigation*[12], the court first required hyperlinked documents to be treated as attachments under the protocol, only to later relax that requirement when defendants showed the burden of compliance.[13] These inconsistent rulings in the US illustrate both the unsettled legal landscape in different jurisdictions and the risks of selective disclosure.

Among these US decisions, we expect the Uber decision to be in line with upcoming Canadian court decisions in the area. The US decisions often turn on what was agreed to in a protocol or what was specifically requested by opposing counsel, as it is a request focused discovery process. In contrast, in Canadian jurisdictions aside from Quebec, litigants have positive obligations to produce relevant (and, in some jurisdictions, material) documents. Canadian courts are unlikely to overlook the failure to capture hyperlinked documents if those are expected to be relevant to the issues in dispute, especially now that their use has become common practice in communications and technologies exist to properly collect them.

Further, proper collection of hyperlinked documents goes directly to the requirements of system integrity under the CEA, s. 31.5, which contemplates proof of the procedures and practices by which records are created and maintained. Likewise, the CGSB-72.34-2024 standard emphasizes that metadata, family relationships, and audit trails are essential to establishing records integrity. If hyperlinked files are ignored or their contemporaneous versions cannot be shown, the evidentiary foundation required for admissibility is undermined.

In practice, counsel need to inquire about their clients' communication technologies and their use early on so that hyperlinked files can be properly collected.

---

[10] 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.
[11] 2024 WL 1772832 (N.D. Cal. 2024)
[12] 2024 WL 2305604 (N.D. Cal. 2024)
[13] StubHub showed that, despite considerable effort (hundreds of hours plus hiring an outside e-discovery vendor), it was unable to locate many of the linked/hyperlinked documents. Most of the hyperlinks were broken or no longer worked.

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 authenticity and the business records exception. These principles were not designed with modern evidentiary realities in mind. The Horizon scandal is a cautionary warning: misplaced confidence in system reliability can distort justice for years before the error becomes visible.

In Canada, the CEA offers presumptions that promote efficiency, but those same presumptions can create imbalance when only one party controls the underlying system. The CGSB standard offers a vocabulary for integrity and reliability, but tribunals must decide whether adopting such frameworks enhances arbitral legitimacy or risks importing rigid formalism into a process valued for its flexibility. Similarly, AI-driven expert analysis may offer speed and scale, but it poses the risk of conclusions that cannot be interrogated. Hyperlinked attachments can now be properly collected, but unless parties address them explicitly, tribunals may be left with evidentiary gaps that affect both completeness and confidence in the record.

None of these are 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 does 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 faced with digital records whose reliability is credibly contested, do you see your role as managing proportionality or as ensuring that the award rests only on reasonably challenged evidence that has been properly proven?
- As AI-generated analysis and system-level evidence become routine, should tribunals 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: Missing, alterable, or incomplete logs make it impossible to reconstruct events, showing poor control over the system.
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. Record Integrity

Record Integrity Arguments (individual records)
Altered content: Evidence that the record has been modified, corrupted, or otherwise changed from its original state defeats record integrity.
Missing or inconsistent metadata: If creation dates, authorship, or version history are incomplete or contradictory, authenticity and identity are undermined.
Multiple versions without version control: Absence of documented versioning procedures raises doubt about which copy is authoritative.
Breaks in chain of custody: Inability to show how the record was captured, stored, transferred, or preserved creates doubt about whether it is complete and unaltered.
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: If a record or its supporting metadata was deleted or destroyed prematurely, its integrity cannot be demonstrated.

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?[14] – Does one party have exclusive access to proprietary systems?
Legal & Ethical Standards	• Compliance with statutory frameworks • Case law precedent • Professional responsibility	– Does the evidence align with Canada Evidence Act s.31.5 and provincial rules? – How have courts treated AI/digital evidence to date? – Has counsel met duties of competence and candour?

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