http://doi.org/10.7207/twgn26-01

Cyber Security and Resilience for Digital Preservation

Part 1: Understand Cyber Threats

Heather Lowrie and Garth Stewart

DPC Technology Watch Guidance Note

June 2026

DigitalPreservationCoalition

This report is published by the Digital Preservation Coalition (DPC). The DPC is an international charitable foundation which supports digital preservation and helps its members around the world to deliver resilient long-term access to digital content and services. In addition to the publication of reports on a range of themes which cover the state of the art in digital preservation, the DPC also supports its members through community engagement, targeted advocacy work, training and workforce development, identification of good practice and standards, and through good management and governance. Its vision is a secure digital legacy.

Discover the DPC's publications, including the latest updates and revisions, at:
https://www.dpconline.org/digipres/discover-good-practice/tech-watch-reports

Find out more about the DPC, the support it offers and how to become a member at:
https://www.dpconline.org/about/join-us

DPC Cyber Security and Resilience Guidance Note Series

Organizations that carry out digital preservation activities operate in an increasingly hostile cyber environment. Threats have grown in frequency, scale, and sophistication, placing unique and irreplaceable collections at risk. Sadly, it is now a case of - *not if, but when* - any organisation holding digital collections of long-term value is subject to cybercrime.

This creates pressure on digital preservation practitioners to understand this issue to a reasonable level of detail and identify key cyber threats to their collections. It is rare that practitioners will be able to confront these threats alone: collaboration and engagement with other teams such as IT will likely be a key next step. This collaboration should lead to the application of appropriate cyber security measures that protect digital collections as far as possible and enable response to a cyber-attack.

However, cyber security is a challenging, often overwhelming subject, especially for practitioners who may have limited time and resources yet need to understand the following areas:

1. Where to begin with a subject like this: what are the risks?
2. What successful pathways towards engagement look like.
3. How to put plans into action so that attacks can be neutralised, impacts can be mitigated, and recovery is possible.

The DPC has published a set of three Guidance Notes to support practitioners in this effort. They will assist practitioners and organizational leaders in developing effective responses while maintaining business continuity and safeguarding staff wellbeing. They also encourage organizations to embed cyber security within digital preservation governance and operational practice in ways that protect collections while enabling usability, access, and long-term resilience.

This Guidance Note focuses on the first part of this approach: understanding cyber threats to digital collections. Note that technical terminology is defined within the Glossary.

1 The Digital Preservation Security Challenge

Digital preservation is all about risk management: a practice intended to identify, analyse and mitigate risks to the long-term accessibility of digital materials. Cybercrime now represents one of the most significant of those risks.

With year-on-year increases in frequency, sophistication, and capability, cybercrime has moved from a peripheral IT concern to a major global economic risk. CEOs and national leaders now rank cyber-enabled threats alongside geopolitical instability and systemic economic disruption. (World Economic Forum, 2026). In the UK alone, independent government research estimates that cyber-attacks cost the UK economy around **£14.7 billion per year**, which is roughly **0.5 % of UK GDP** (UK Government, 2025).

What does this mean for digital preservation? A complex security challenge. Digital preservation systems maintained by libraries, archives and preservation repositories across the World contain unique, high value, and sometimes sensitive collections. The significance of these collections cannot be overstated. In a lightning talk at iPRES 2025, Yannick Grandcolas from Bibliotheque nationale de France described his institution as, 'more than just a library, it is a guardian of France's cultural and intellectual heritage' (Grandcolas, 2025). Such value, depressingly, creates a bigger target for cybercrime.

Furthermore, collections can often be managed by relatively small teams operating under significant resource constraints. This combination of value, technical complexity, and constrained resourcing make for attractive targets for cyber criminals. It can also create security challenges that differ substantially from those faced in conventional enterprise IT environments.

Inevitably, in recent years this has led to many cyber-attacks on organizations doing digital preservation. No sector has remained untouched: higher education, healthcare, local government, and even national institutions have all been attacked, including the severe 2023 cyber-attack on the British Library (British Library, 2024).

The consequences of attacks and associated data breaches can be devastating. The authenticity, integrity, availability and confidentiality of collection data may be compromised, stolen, corrupted, or permanently lost. Organizations can incur substantial financial costs. Operations may stall or cease entirely. Reputational damage and legal liabilities may emerge. On a human level, attacks can have severe negative impact on the health and wellbeing of staff (British Library, 2026).

This worsening situation led William Kilbride, Executive Director of the DPC in 2023 to call for a 'blue shield' for digital cultural heritage to be formed, so that it may be 'off limits to cyberwarfare' (Kilbride, 2023). Cybercrime features prominently in the Executive Summary of the 2025 Global 'Bit List' (Digital Preservation Coalition, 2025).

But what practical steps can digital preservation practitioners take to respond to these threats? Before effective mitigations can be developed, it is necessary to understand the cyber risks to which digital preservation environments are especially exposed.

2 Understanding Cyber Threats and Risks

2.1 Global Threats

On a general level, the cyber threat landscape includes both conventional attacks and emerging risks to digital preservation activities. Key threats include:

- **Malware**: software designed to damage computer systems, networks or devices. Computer viruses are one form of malware which pose challenges to the digital preservation community.
- **Ransomware:** a particularly destructive form of malware. If deployed within a repository, ransomware can block user access while encrypting data, with criminals demanding payment to restore access. Payment does not guarantee recovery, and data may still be stolen, destroyed, or publicly leaked. Increasingly, ransomware campaigns are targeting backups and archives directly, bypassing traditional endpoints such as user laptops and servers, thereby threatening the recoverability of irreplaceable digital collections.
- Organizations frequently rely on third-party technologies, cloud providers, and software dependencies procured through complex supply chains 'as a service', for example Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS). While these relationships enable scalable and cost-effective technical infrastructures, they can create opportunities for **supply-chain compromise**, whereby trusted tools, vendors, or services are infiltrated and used to introduce malicious code, vulnerabilities, or unauthorised access into preservation environments.
- Like conventional IT systems, account management within digital preservation environments remains vulnerable to **phishing and spear-phishing** attacks, where malicious actors deceive users into disclosing credentials or granting system access through fraudulent communications and harmful links.
- Recent advances in **quantum computing** also present potential risks to digital preservation. Future quantum capabilities may compromise current cryptographic methods, including checksum and hashing algorithms, thereby threatening the long-term confidentiality and integrity of preserved content.
- The evolving role of **Artificial Intelligence (AI)** currently introduces two core security challenges:
    - The so-called 'post-human future', whereby highly sophisticated AI-generated phishing and deepfakes are becoming standard social engineering tools, and growing cyber threats (Threatdown, 2026). These could pose new threats to authenticity and evidential trust placed in collections, even after ingest.
    - AI-driven scraping of the Web, enabling large-scale unauthorised harvesting of repository content for commercial or research purposes, is increasing the likelihood of distributed denial of service (DDoS) attacks and associated impacts:
        - Unauthorised reuse: Content may be harvested without permission, potentially infringing legal agreements or donor conditions.
        - Disclosure of sensitive information: Large-scale aggregation may inadvertently expose restricted or sensitive material.
        - Service impact: High-volume automated scraping can degrade system performance or disrupt access for legitimate users.
        - Loss of contextual integrity: Aggregated data can be misinterpreted if removed from its original archival context (Weinberg, 2025).

2.2 Particular risks with Digital Preservation

As a practice, digital preservation processes can introduce risks and operational requirements that are often poorly accounted for within conventional cyber security models. These risks can be grouped into three categories:

2.2.1 Content Risks

- Ingest of collections with **unknown or mixed provenance**, which may include malicious or corrupted files. Content can contain embedded executable code or scripts, which can trigger malware alerts or exploit vulnerabilities.
- Content being preserved can contain sensitive personal or organizational data requiring strict privacy and access controls. As well as the content itself, the measures used to enforce these controls must persist for decades, even hundreds of years. This **requirement for perpetual retention and access controls** distinguishes digital preservation systems from mainstream IT environments, which typically assume much shorter content lifecycles.
- Requisite **large-scale processing of diverse file formats** increases exposure to risk through high-volume automated workflows. While automation can mitigate these risks, implementing robust solutions may be complex and costly. Conversely, manual processing introduces its own risks, including human error and missed threat detection.

2.2.2 Technical Complexity

- Digital content can require **legacy** operating systems and obsolete software for access, that cannot be easily patched, secured or rebuilt.
- The use of virtualised and containerized **processing environments** for key preservation tasks such as content receipt and digital forensics introduces additional dependency, configuration, and supply chain risks.
- Historical and continued dependence on **open-source** tools, sometimes with limited security review or support can introduce risks.
- For reasons of cost efficiency and scalability, there is increasing dependence on **cloud services** for storage and processing. While this can improve resilience against data loss, it also expands the overall 'attack surface' and increases exposure to security vulnerabilities.
- Opaque web-facing services and **AI-enabled** components can be embedded into preservation workflows, even though their decision making and security properties may not always be fully understood or auditable. If deployed in proximity to where collections are kept and being preserved, results could be disastrous.

2.2.3 Operational Requirements

- Routine ingest of content extracted from **external media**, including USB drives, hard drives, optical discs, and other legacy carriers is a common requirement. Connecting such media to hardware is often prohibited in modern organizations, as malware can easily leak through this route (CISA, 2021).
- Digital preservation practice is continually evolving. Experimentation with **new tools and file formats,** often necessary for handling rare or unusual content, can create potential vulnerabilities and entry-points.

- There is a need to support diverse internal and external users, each with differing **access** needs and security expectations.

2.3 The Supply Chain

Finally, digital preservation increasingly relies on interconnected ecosystems of software, hardware, cloud services, vendor products, and open-source technologies. There are many ways to procure technical solutions, often in modular, affordable form such as Software-as-a-Service (for example, preservation software), or Infrastructure-as-a-Storage (for example, Cloud storage). These can be supplemented by using open-source tools for particular tasks (for example, full text or metadata extraction), towards preservation infrastructures that exist through associated supply chains.

Supply chains are useful as they enable organizations to build and scale digital preservation capability quickly. However, they also create their own cyber security vulnerabilities. Any supply chain component may become an entry point for malicious activity across the wider infrastructure.

2.3.1 Key Risks

- **Compromised dependencies or deliberate malicious intent**: Trusted software libraries or tools can be altered to include vulnerabilities or malicious functionality, either directly or through upstream compromise.
- Commercial vendor **failure or lock-in**: Reliance on a single provider for all aspects of digital preservation can create severe operational risk if the vendor goes out of business, fails to deliver security updates, restricts data access, or does not deliver on their statements of security compliance.
- **Data sovereignty** and contractual limitations: Cloud or outsourced services may store data in jurisdictions with different legal protections, affecting long-term preservation obligations.
- **Service disruptions** or contractual changes: Cloud outages, deprecations, or modifications in terms of service can interrupt preservation workflows. Even major services like Microsoft Azure are not immune, and cloud data centres are targeted in modern warfare (Speed, 2026, Jarnecki and Sylvia, 2026).
- **Open-source project risks**: Community-developed software may be under-resourced, inadequately reviewed, or ultimately abandoned, leaving vulnerabilities unaddressed.

When using supply chain components in their work, digital preservation practitioners are encouraged to consider these vulnerabilities in their risk assessment.

3 Conclusion

The threats and risks in this section reflects the state of knowledge as of 2026, and updates are available via authoritative sources (See section 5.1).

Each of these risks can directly threaten the long-term accessibility, authenticity, and integrity of digital collections if not actively managed: especially if they occur in combination. It is important for digital preservation practitioners to understand and evaluate these risks in according with their organization's risk appetite. In general, a strong security-first or low-tolerance approach to cyber threats is recommended, and the DPC's '*A Risk Driven Approach to Bitstream Preservation*' may be useful in considering this (Wheatley, 2025).

Digital preservation practitioners do not face these challenges alone. Cybersecurity is a challenge which must be tackled collaboratively with other teams within an organization. Articulating and

explaining to related stakeholders - such as IT - the needs and challenges of digital preservation and how these can fit into and indeed strengthen cybersecurity operations, is a key task. The topic of engagement and communication forms the basis of the next Guidance Note in this series.

4 Glossary

Attack surface: The sum of vulnerabilities, pathways, or methods—sometimes called attack vectors—that hackers can use to gain unauthorized access to an organization's network or sensitive data, or to carry out a cyber-attack.

Authenticity: That information contained within digital content remains trustworthy.

Availability: That information contained within digital content can be located and rendered.

Confidentiality: That information contained with digital content is stored and accessed securely.

DDoS (Distributed Denial of Service): A cyber-attack in which multiple compromised systems flood a target (such as a website or server) with traffic, overwhelming it and making it unavailable to legitimate users.

Deepfake content: AI generated fake content such as images, videos and audio designed to deceive. These materials are often being used by criminals to deceptively steal money, strip away the dignity of women and girls, and spread harmful content online.

Encryption: A security process that converts readable data (plaintext) into unreadable code (ciphertext) using algorithms and keys, so that only authorized parties can access the original information.

Infrastructure-as-a-Service (IaaS): A service model that allows customers to access virtualised hardware and software resources over the internet, often via the cloud.

Integrity: That information contained within digital content remains accurate and complete.

Malware: Malicious software designed to disrupt, damage, or gain unauthorized access to systems. Examples include viruses, worms, ransomware, spyware, and trojans.

Phishing: Where attacks send scam messages to users which contain malicious links in order to spread malware or gain entry to sensitive information.

Platform-as-a-Service (PaaS): A Cloud-computing service model that allows customers to build and deploy applications upon an underlying platform provided and managed by a cloud provider.

Ransomware: A type of malware which prevents access to devices and data stored on these, usually by encrypting files. A criminal group will then demand a ransom in exchange for decryption.

Scraping: The act of finding, downloading, and formatting data from the internet via automated means (for example via web crawl).

Social engineering: Term used for a broad range of malicious activities accomplished through human interactions. It uses psychological manipulation to trick users into making security mistakes or giving away sensitive information.

Software-as-a-Service (SaaS): Cloud-computing service model whereby a cloud provider runs an application in a way that is shared between its customers. Customers only have to consider how to configure and consume the app so that it meets their needs. Common method of deployment for digital preservation supplier solutions.

Spear Phishing: A form of phishing targeting particular individuals where the email is designed to look like it's from a trusted or known person.

5 References

British Library (2025) *Learning lessons from the cyber-attack.* London: BL. Available at: https://web.archive.org/web/20260608000132/https://www.bl.uk/stories/blogs/posts/learning-lessons-from-the-cyber-attack

Center for Internet Security (2024) *Critical Security Controls*. East Greenbush: CIS. Available at: https://www.cisecurity.org/controls/ [Accessed 11 February 2026].

Cybersecurity & Infrastructure Security Agency (2021) *Using caution with USB Drives*. Available at: https://web.archive.org/web/20260615123500/https://www.cisa.gov/news-events/news/using-caution-usb-drives

Digital Preservation Coalition (2025) *Global Bit List of Endangered Digital Materials Report 2025*. Available at: https://web.archive.org/web/20260529141818/https://www.dpconline.org/digipres/champion-digital-preservation/bit-list

European Union Agency for Cybersecurity (2025) *Threat Landscape Reports*. Heraklion: ENISA. Available at: https://www.enisa.europa.eu/publications [Accessed 11 February 2026].

Grandcolas, Y. (2025) *Digital Sovereignty Unveiled: Safeguarding France's Cultural Legacy at the National Library.* iPRES. Available at: https://doi.org/10.5281/zenodo.19447287 [Accessed 15 June 2026].

International Organization for Standardization / International Electrotechnical Commission (ISO/IEC), 2013. *ISO/IEC 27001: Information Security Management Systems – Requirements*. Geneva: ISO/IEC.

Jarnecki, J. and Sylvia, N. (2026) 'Iran Data Strikes Shae Global Digital Infrastructure', *Royal United Services Institute (RUSI).* Available at: https://web.archive.org/web/20260324222449/https://www.rusi.org/explore-our-research/publications/commentary/iranian-data-strikes-shake-global-digital-infrastructure

Kilbride, W. (2023) *DP Disrupted: A blue shield for the electronic battlefield*. Available at: https://web.archive.org/web/20260615130027/https://www.dpconline.org/blog/dp-disrupted-a-blue-shield-for-the-electronic-battlefield

Open Web Application Security Project (OWASP) (2023) *Software Supply Chain Security Guidance*. Available at: https://web.archive.org/web/20260412195218/https://cheatsheetseries.owasp.org/cheatsheets/Software_Supply_Chain_Security_Cheat_Sheet.html

Speed, R. (2026) 'Azure outages ripple across multiple dependent Microsoft services', *The Register.* 3 February 2026. Available at: https://web.archive.org/web/20260204005424/https://www.theregister.com/2026/02/03/azure_virtual_machine_outage/?td=rt-3a

The Times (2026) 'British Library boss: We're still healing emotionally from 2023 hack', *The Times*. Available at: https://www.thetimes.com/uk/crime/article/british-library-hacking-cyberattack-rqrgjch99 [Accessed 25 February 2026]

Threatdown, (2026) *2026 State of Malware*. Available at: https://www.threatdown.com/dl-state-of-malware-2026/ [Accessed 26 February 2026]

UK Government, 2025. *Cyber Security Breaches Survey*. London: Government Digital Service. Available at: https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2025 [Accessed 11 February 2026].

Weinberg, M. (2025) *Are AI Bots Knocking Cultural Heritage Offline?*, Exeter: GLAM-E Lab Authors. Available at: https://web.archive.org/web/20260507135729/https://www.glamelab.org/products/are-ai-bots-knocking-cultural-heritage-offline/

Wheatley, P. (2025) *A Risk Driven Approach to Bitstream Preservation*. Digital Preservation Coalition. Available at: https://web.archive.org/web/20260309073104/https://www.dpconline.org/docs/technology-watch-reports/2793-bitstreampreservationrisk/file

World Economic Forum (WEF (2025) *Global Risks Report 2025*. Geneva: WEF. Available at: https://www.weforum.org/reports/global-risks-report-2025/ [Accessed 11 February 2026].

5.1 Useful websites

The cyber threat landscape, and good practice to on how to mitigate it, are in a constant state of flux. It is useful to refer to authoritative cyber security agency advice for the latest intelligence on this. A few key websites are listed below, and practitioners are encouraged to seek out similar sources of advice applicable to their own region if not listed here.

Agence nationale de la sécurité des systèmes d'information (France): https://cyber.gouv.fr/

Australian Signals Directorate: Cyber Security Centre: https://www.cyber.gov.au/

Cybersecurity & Infrastructure Security Agency (USA): https://www.cisa.gov/

European Union Agency for Cybersecurity (EU): https://www.enisa.europa.eu/

National Cyber Security Centre (UK): https://www.ncsc.gov.uk/

Websites all accessed 11 June 2026.