

# Affordable digital preservation for libraries and museums

Ensuring your digital content and collections  
are safe and accessible for future generations







# Executive summary

Most libraries, archives, museums and other cultural heritage institutions – big and small - are facing the challenge of how best to safeguard and make accessible the rapidly growing volume and diversity of digital content – either from digitization programs or through the acquisition of original born-digital material.

As highlighted in the recent POWRR report for the Institute of Museum and Library Services<sup>1</sup> digital content is inherently fragile, at risk due to a number of factors, including basic data corruption, storage media degradation and obsolescence, and file format and software program obsolescence.

As custodians of often unique digital content it has become important for the modern library and museum to have a robust strategy, processes and systems in place for handling, managing and providing access to digital material. Sound digital preservation practices in many cases have become fundamental to attracting and properly serving donors, patrons, researchers and the public.

This guide explores these challenges in more detail, highlighting how other institutions have tackled the challenge and proposing a proven, standards based approach<sup>2</sup> to assessing digital preservation capability and selecting a suitable and affordable digital preservation solution.

The aim is to help you gain a clearer picture of what you need to do in order to ensure your unique digital collections and archives are properly safeguarded, accessible and more importantly still usable for generations to come.

---

1. From Theory to Action: "Good Enough" Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions. A digital POWRR White Paper for the Institute of Museum and Library Services. Full report available via: <http://preservica.com/resource/ims-powrr-report/>

2. Assessing Digital Preservation Capability Using a Maturity Model Process: White Paper Charles M Dollar and Lori J. Ashley <http://www.savingthedigitalworld.com/papers-research>

2a. ISO 14721 Open Archival Information System (OAIS) Reference Model [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=57284](http://www.iso.org/iso/catalogue_detail.htm?csnumber=57284)

2b. ISO 16363 Audit and Certification of a Trustworthy Digital Repository [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=56510](http://www.iso.org/iso/catalogue_detail.htm?csnumber=56510)

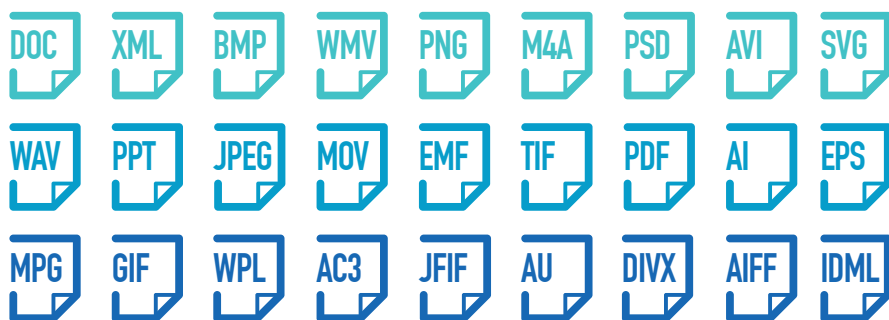


# Contents

1. Why digital content and collections need special care
2. The key elements of a robust digital preservation solution
3. Selecting a complete standards based digital preservation solution
4. Understanding the opportunity of the Cloud
5. Combining Digital Preservation and Public Access into one system
6. How other institutions have tackled the digital preservation challenge
7. Integration with other systems
8. Useful resources



# Why digital content and collections need special care



As highlighted in the recent POWRR report for the Institute of Museum and Library Services (IMLS)<sup>1</sup> - both research conclusions and professional's real-life experiences agree that digital content is fragile – susceptible to bit corruption, media degradation (have you checked your CD-ROMs recently?), technology obsolescence (remember floppy disks?) and file format and software obsolescence (remember WordStar?).

Compounding the challenge is the rapidly increasing volume and diversity of digital material that institutions are being asked to curate and care for - both from digitization programs and the acquisition of born-digital content.

All this means that digital collections and archives need special care to ensure that these unique records are properly preserved, accessible and more importantly still readable and usable by future generations.

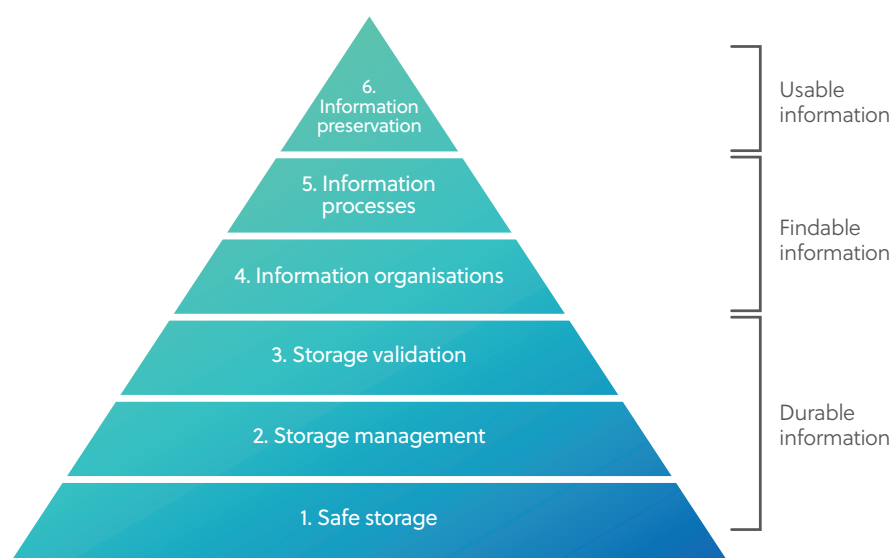
This is where a standards-based digital preservation solution comes in - that provides a top to bottom approach to storing, managing and preserving digital content.

And as we will explore later – simply storing and making your collections accessible through a digital library service is not sufficient to properly safeguard your valuable digital content.

---

1. From Theory to Action: "Good Enough" Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions. A digital POWRR White Paper for the Institute of Museum and Library Services. Full report available via: <http://preservica.com/resource/imls-powrr-report/>

# The key elements of a robust digital preservation solution



The Active Preservation of Digital Content © Preservica 2014<sup>1</sup>

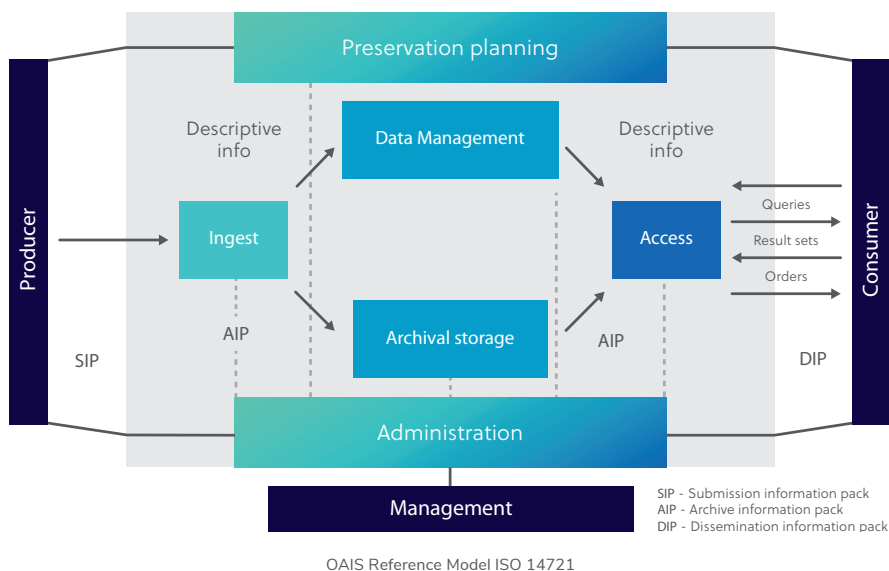
Curating and caring for digital content requires a strategy that encompasses policy, processes, governance, people and systems. This is explored more fully in a recent White Paper which outlines a Digital Preservation Capability Maturity Model (DPCMM)<sup>2</sup> for assessing an organizations digital preservation capability across all these functions with view to achieving an ISO standard trusted digital repository (ISO 16363).

A core element in this strategy is the role played by the digital preservation system itself. As outlined in the diagram above, any system needs to go beyond simply storing the bits and bytes in a durable way – such as multiple copies, multiple locations, integrity checking and self-healing (Level 1, 2 & 3 in the diagram) to include information organization, management, hierarchies, workflows and metadata (Level 4 & 5) so content can be easily found and understood in the future, and finally – and most importantly – it needs to provide a way to proactively manage and migrate file formats as they become obsolete to ensure valuable digital collections are actually still usable and readable by future generations (Level 6).

1. For more detail on the Active Preservation of Digital Content - download the Preservica White Paper: <http://preservica.com/resource/auctor-aliquam-white-paper/>

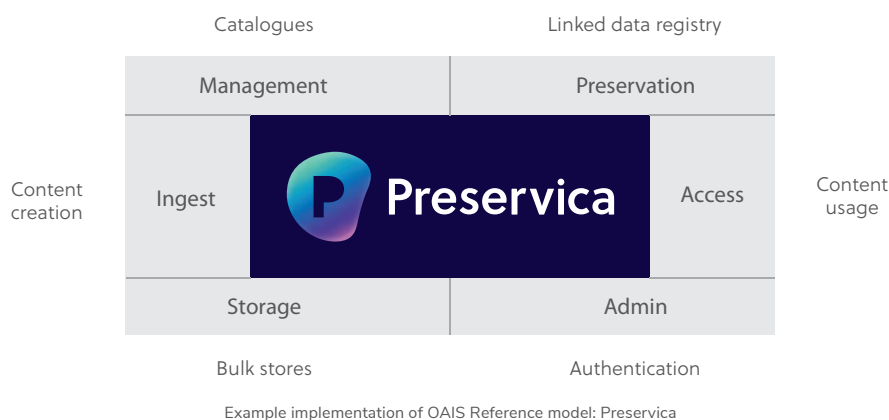
2. Assessing Digital Preservation Capability Using a Maturity Model Process: White Paper Charles M Dollar and Lori J. Ashley <http://www.savingthedigitalworld.com/papers-research>

# Selecting a complete standards based digital preservation solution



A complete digital preservation system should also conform to the main standard for archiving and preservation, namely ISO 14721 the Open Archival Information Systems (OAIS) Reference model.

This means you need to look for an integrated solution that has workflows and capabilities across all OAIS functions – ingest, data management, storage, administration, access – and especially preservation planning to ensure that content can be easily migrated to newer file formats as old formats become obsolete.





Copy	Fixity check	Virus scan	File dedupe	Auto unique ID	Auto metadata creation	Auto metadata harvest	Manual metadata	Rights management	Package metadata	Auto SIP creation	Public interface	Auto DIP creation	Auto AIP creation	Reliable, long-term bit protection	Redundancy	Geographically dispersed data storage model	Exit strategy	Migration	Monitoring	Auto recovery	
Ingest					Processing						Access		Storage				Maintenance				
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

POWRR Tool Grid showing how Preservica Cloud Edition addresses all 21 functions

The OAIS Reference model, along with the Digital Curation Centre's Digital Duration Lifecycle<sup>1</sup>, were also recently used by the POWRR<sup>2</sup> team in their report for the Institute of Museum and Library Services, to create their Grid for measuring the completeness of different digital preservation tools and solutions against 21 functions.

1. DCC Curation Lifecycle Model <http://www.dcc.ac.uk/resources/curation-lifecycle-model>

2. From Theory to Action: "Good Enough" Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions. A digital POWRR White Paper for the Institute of Museum and Library Services. Full report available via: <http://preservica.com/resource/imls-powrr-report/>





# Understanding the opportunity of the Cloud

Cloud storage services, such as Amazon Web Services S3 and Glacier, are ideal for the preservation of both digitized and born-digital content.

They are inherently durable (multiple copies, multiple locations, integrity checking and self-healing) and also relatively low-cost (with regular price reductions) as well as secure and highly scalable.

Different services can be used for different types of content e.g. presentation and access copies held in a low-latency service like Amazon S3, and infrequently accessed high-resolution digitization images in a low cost service like Amazon Glacier.

However, as discussed in Section 2, when considering Cloud storage for digital preservation make sure you look for a solution that includes workflows and functions that ensure your digital assets are easily findable - and more importantly still usable and readable long into the future. Also, unless you are building a “dark archive” then rapid internal and public access to your digital collections needs to be considered.

## Cloud hosted digital preservation

If you are a small to mid-size institution with limited in-house technical resources then look for a digital preservation and access solution that is hosted in the Cloud. This type of solution should provide you with everything you need right-of-the-box enabling you to very quickly start safeguarding your valuable digital content. And because everything is fully hosted and maintained for you then more time can be spent curating your digital collections - rather than managing and supporting local tools, servers and storage systems. This makes Cloud hosted digital preservation a very cost-effective and affordable choice.

## On premise digital preservation with Cloud storage

Larger institutions with more technical resource might consider an on premise digital preservation and access solution. In this scenario, Cloud storage can also be used in conjunction with on premise storage to create a hybrid model that provides additional scalable redundancy and durability in a very cost effective way.

Finally, there are a number of considerations to choosing the Cloud for digital preservation – from security to having an exit strategy. In addition, any digital preservation solution cloud hosted or on premise should offer flexibility over where and how your valuable content gets stored.

These considerations are explored in more detail in a recent Preservica webinar<sup>1</sup> as well as in various Amazon Web Services' White Papers<sup>2</sup>

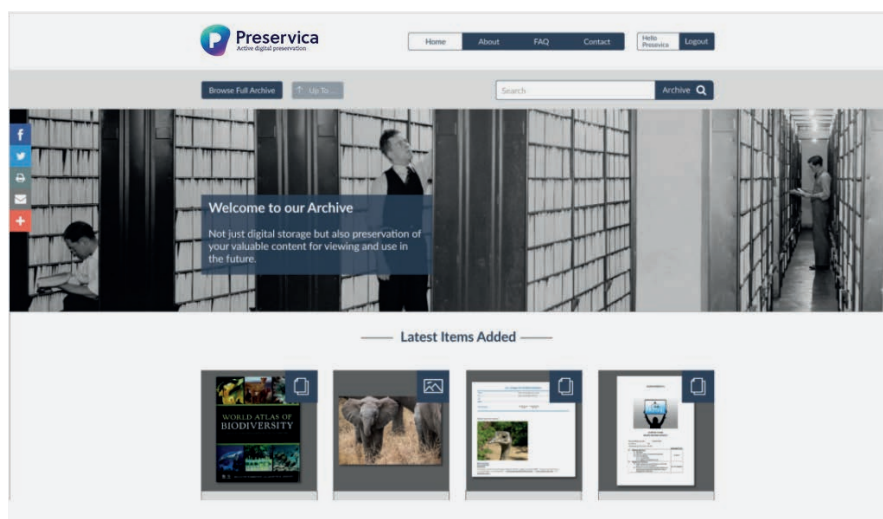
---

1. Digital Preservation in the Cloud webinar <http://preservica.com/resource/awswebinar/>

2. Amazon Web Services Cloud White Papers [http://aws.amazon.com/whitepapers/?nc1=f\\_cc](http://aws.amazon.com/whitepapers/?nc1=f_cc)



# Combining digital preservation and access into a single system



Preservica Universal Access - showing customized Public Access view of digital collection<sup>1</sup>

Providing and expanding public access to content and collections is a fundamental requirement for most institutions - and a key foundation to securing long term sustainability.

Today, online access is often provided via a digital library platform, such as CONTENTdm.

However, these solutions are more concerned with providing online access to presentation versions of digital collections - rather than full digital preservation e.g. of high-resolution digitized images or the original format like a Word document.

Many forward thinking institutions such as Grand Valley State University Libraries<sup>2</sup> have realized that by choosing a solution that combines digital preservation and public access in one they can move away from having a separate digital library to consolidating preservation and access into one system – saving them time, money and duplicated effort.

In addition, a modern public access solution also needs to ensure your collections and archives are easy to find, browse, search and view using a variety of different devices including PCs, Tablets and Smart Phones. You should also be able to securely control which content is made public, and which is not, as well as easily customise your public site so it stands out from the crowd and reflects the brand values of your institution.

1. Preservica Universal Access <http://preservica.com/public-access-discovery/>

2. To learn more about GVSU Libraries' use of digital preservation and plans you can watch a recording of this recent Preservica webinar <http://preservica.com/resource/contentdmwebinar/>



# How other institutions have tackled the digital preservation challenge



**Background:** DC Public Library is a leading US Public Library with an extensive and varied multi-media digital collection including: Clifford Berryman cartoons from the Washington Evening Star 1900 to 1948; photographs by Joseph Owen Curtis documenting the culture, social life, and architecture of Southwest Washington D.C. from the 1920s through the 1970s; oral histories of first-hand accounts of the historic March on Washington for Jobs and Freedom in 1963; and music, videos, images and other content from the DC Punk movement of the 1970s, 80s and 90s. <http://digdc.dclibrary.org/>

**Challenge:** Like many institutions, DC Public Library was faced with a growing volume and diversity of born-digital content as well as content from their own digitization program. This prompted the special collections team to commission an independent study to outline what they needed to do to properly safeguard and make their unique digital content accessible for future generations.

As well as making recommendations on governance and people, a key recommendation of the study was to ensure the team deployed a standards-based Digital Preservation system to securely store and preserve digital content. As a relatively small department it was also important that any system was fully hosted, managed and maintained in the Cloud in order to free up the team's time to focus on curating their collections and not worry about integrating and maintaining individual tools or buying and operating local servers and storage.

**Solution:** DC Public Library appointed a Digital Curation Librarian, who set about assessing different tools and systems against the departments needs and the core functions of the OAIS (Open Archival Information System) Reference model (in a similar fashion to the POWRR Tool Grid). This meant that the chosen system not only needed to be able to store digital content in a secure, durable way, but also make it easy to manage and migrate obsolete file formats to newer formats overtime.

**Future:** In the future the special collections team plan to explore using a single system for both digital preservation and public access in order to migrate away from their digital library platform. In addition, they would like to utilise the flexibility of their chosen system to exchange metadata with the Digital Public Library of America.



**Background:** Grand Valley State is a leading Michigan, US based liberal arts University. The library's extensive and varied digital multimedia collection, used for research and study, includes diaries and correspondence from the Civil War era, slavery documents, 15th century printing; photographs and clippings from American military veterans from World War I to the Afghanistan and Iraq conflicts; interviews and papers of the Young Lords in Lincoln Park, 1960s community activists; oral histories of Grand Rapids area founders, philanthropists, and interfaith leaders; D.J. Angus photographs and moving pictures of ca. 1930s US and Mexico; and other highlights of the holdings in the library's Special Collections and University Archives. <http://gvsu.cdmhost.com/cdm>



**Challenge:** The library's initial focus was on making their Collections and University Archives available online to students and the public using a digital library platform. But with over 10 TBs of original digital material (growing at 1 – 2 TB per year) the team also realized they needed a secure and durable way to properly preserve this unique digital content for future generations.

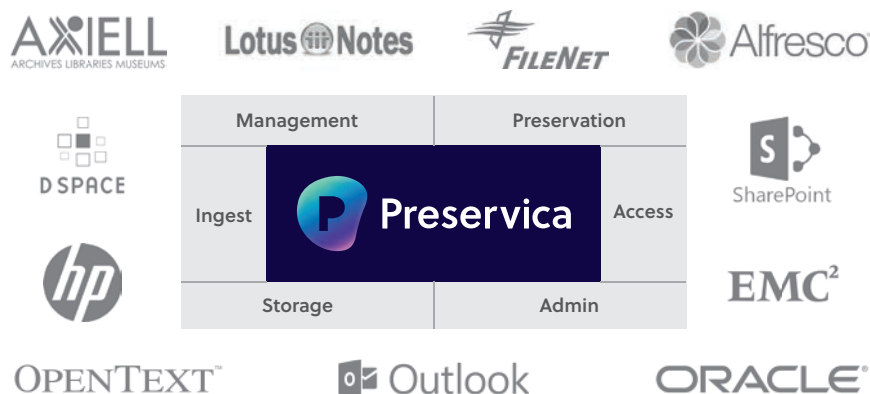
In particular, as well as providing a secure storage environment any system would also need to make it easy to manage and migrate obsolete file formats to newer formats overtime, while maintaining metadata across the object's lifecycle. The system also needed to be fully Cloud hosted, able to provide all the digital preservation functions the team needed right out-of-the box. This would free up time to spend on curation – and not on maintaining individual tools or managing local servers and storage.

**Solution:** The Metadata and Digital Curation Librarian was tasked with evaluating the available digital preservation solutions. Preservica Cloud Edition was chosen – because of its completeness against the OAIS reference model functions, the fact that it is fully Cloud hosted, and that it was a fraction of the cost of its nearest on premise only rival.

**Future:** The team were also impressed that Preservica Cloud Edition was able to provide full digital preservation and public access in a single system (at an affordable price) and are making plans to migrate away from their online digital library service and consolidate into one system. This will enable them to save cost and simplify the way they manage and publish digital content going forward.



## Integration with other systems



Your institution will probably have a wide variety of different information systems e.g. for collection and archive management, library management and cataloguing, as well as systems for operational records, such as a content/ records management system. Any digital preservation solution you are looking at should therefore easily connect into your overall information landscape.

### Hybrid Archives and Collections

Most institutions have hybrid archives and collections – a mix of physical, digitized and born-digital materials. So ensure your chosen solution either already includes integration and synchronization with a catalogue system (e.g. Axiell CALM, Adlib) or uses standard APIs like OAI-PMH to make connection possible.

Learn how Wellcome Library, one of the world's largest resources for the study of medical history (with over 5M digitized images) is tackling the challenge of having a large scale hybrid archive <http://preservica.com/resource/hybridarchives/>

### Digital Libraries

As we have already explored (in section 5) how consolidating digital preservation and public access into a single system can save money and duplicated effort. Ensure your chosen digital preservation solution includes automated bulk ingest connectors for the common digital library platforms e.g. DSpace and CONTENTdm.

### Operational Records

Digital preservation can also be used to safeguard the long term and permanent operational records of your institution e.g. legal and contractual documents and emails. Using the system for more than just digital collections can help you build an even stronger business case. Ensure your chosen digital preservation solution either already includes connectors for common Enterprise Email and Content/ Records Management systems like Microsoft Outlook and Microsoft SharePoint and/or there is a clear roadmap for providing connectors to other platforms.

### Metadata Aggregation

If it is important for your content to be found, featured and available to services like the Digital Public Library of America (DPLA), then look for a digital preservation solution that has the flexibility to be easily expanded to share your metadata and content.



## Useful resources

From Theory to Action: “Good Enough” Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions. A digital POWRR White Paper for the Institute of Museum and Library Services. Full report available via:

<http://preservica.com/resource/imls-powrr-report/>

Assessing Digital Preservation Capability Using a Maturity Model Process: White Paper Charles M Dollar and Lori J. Ashley

<http://www.savingthedigitalworld.com/papers-research>

ISO 14721 Open Archival Information System (OAIS) Reference Model

[http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=57284](http://www.iso.org/iso/catalogue_detail.htm?csnumber=57284)

ISO 16363 Audit and Certification of a Trustworthy Digital Repository

[http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=56510](http://www.iso.org/iso/catalogue_detail.htm?csnumber=56510)

DCC Curation Lifecycle Model

<http://www.dcc.ac.uk/resources/curation-lifecycle-model>

The Active Preservation of Digital Content - Preservica White Paper:

<http://preservica.com/resource/auctor-aliquam-white-paper/>

Digital Preservation in the Cloud a Preservica and AWS webinar

<http://preservica.com/resource/awswebinar/>

Amazon Web Services Cloud White Papers

[http://aws.amazon.com/whitepapers/?nc1=f\\_cc](http://aws.amazon.com/whitepapers/?nc1=f_cc)

Preservica cloud hosted and on premise Editions

<http://preservica.com/editions-pricing/>

GVSU Libraries' use of digital preservation and connector to CONTENTdm a Preservica webinar

<http://preservica.com/resource/contentdmwebinar/>

Preservica Universal Access: Public Access and Discovery

<http://preservica.com/publicaccess-discovery/>

Getting started: Preservica 5 Step Digital Preservation Journey

<http://preservica.com/services/>





# About Preservica

Preservica is a world leader in digital preservation technology, consulting and research. Our active preservation solutions are used by leading archives, libraries, museums, government organizations and businesses across 4 continents to safeguard and share their valuable digital content, collections and electronic records for decades to come.

The award-winning Preservica Active Preservation and Access technology is available in cloud hosted and on premise Editions and includes a full suite of OAIS (Open Archival Information System) workflows for ingest, management, storage, access and long-term preservation of digital content.

Preservica invests in many research initiatives in digital preservation and is a well-respected member of many international collaborations with academia, archives, libraries and corporations at the leading edge of this emerging field. Preservica is part of Tessella group, a leading international analytics, software services and consulting company.

[www.preservica.com](http://www.preservica.com)

[info@preservica.com](mailto:info@preservica.com)

[@dPreservation](#)