



**Preserving AI**

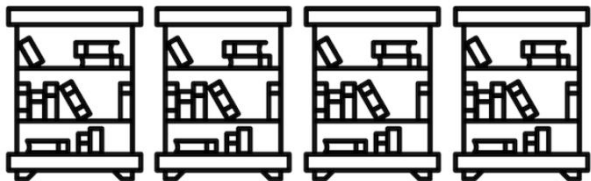
**Precedent, Challenges, and Opportunities**

Thomas Padilla  
Associate Dean, Research and Learning  
University of Nebraska Lincoln

6.5.26



## Always Already Computational

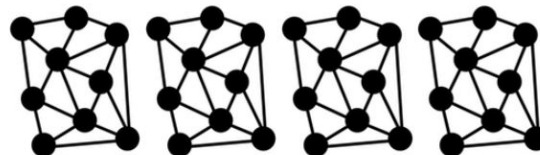


OC LC RESEARCH POSITION PAPER

## Responsible Operations: Data Science, Machine Learning, and AI in Libraries

**Thomas Padilla**  
Practitioner Researcher in Residence

## Part to Whole



## The Public Interest Corpus: A Framework for Implementation

By the Public Interest Corpus Project Team<sup>1</sup>

Introduction And Background: The Need And Opportunity

~2 years ago

INTERNET ARCHIVE



INTERNET ARCHIVE



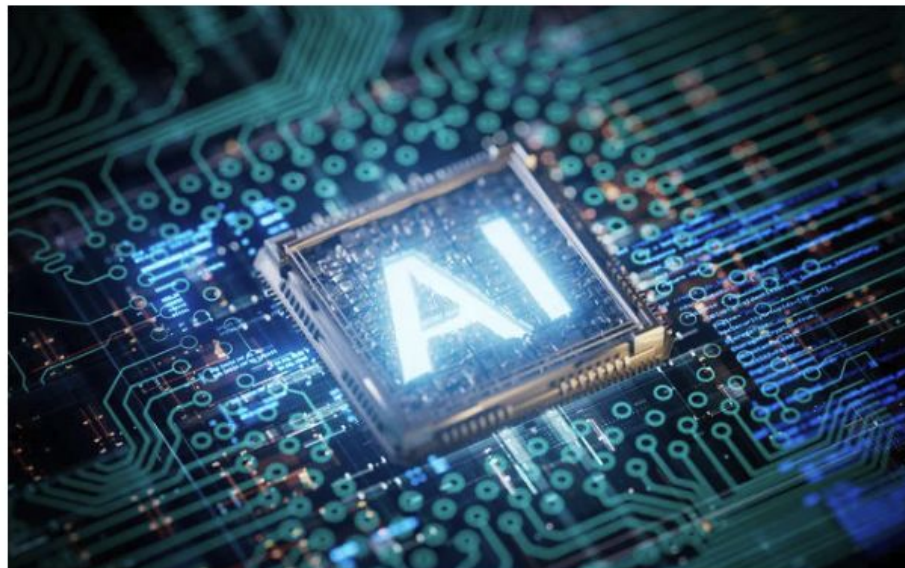
**Hugging Face**

~1 year ago

# Preserving AI

Future scholars will need an enduring digital record of AI technology, Thomas Padilla writes.

By [Thomas Padilla](#)



Advertisement

How can we best understand society moving forward without ongoing access to some version of the tools that fundamentally affect how knowledge in our time is produced?

We have work to do aligning research libraries, AI service providers, philanthropy and policymakers so they can collaborate on long-term AI preservation.

Alignment will be key to resourcing a stable transition from maximally useful AI to an artifactual state available for future study.

AI can and must be preserved.

- 1. precedents**
- 2. challenges**
- 3. opportunities**

**1. precedents**

~~**2. challenges**~~

~~**3. opportunities**~~

# Digital Preservation Theory and Practice

- Collectively, ~ 30 years of LAM experience
- Efforts come and go, but expertise largely persists (for now)
- Robust theoretical and practical experience
  - bit-level preservation
  - emulation
  - significant properties
  - digital forensics
  - custom tooling
  - NDSA levels of preservation
  - centralized and decentralized preservation (e.g., LOCKSS vs. decentralized web tech)
  - Email preservation
  - Software preservation
  - Video game preservation
  - Web archive - and everything therein preservation work

## Also in this section

[What is digital preservation?](#)

## Discover Good Practice

[Get help with your digital preservation](#)[Dedicated Support for Members](#)[DPC Consultancy](#)[Digital Preservation Handbook](#)[Technology Watch Publications](#)[Digital Preservation for Small Businesses: An Introductory Guide](#)ENHANCED BY [Home](#) > [Digital Preservation](#) > [Discover Good Practice](#) > [Get help with your digital preservation](#)

# Get help with your digital preservation

*Use this to find the assistance you need to do digital preservation, from reference materials to asking the question of others*

There are an almost endless number of challenges in the digital preservation field. Whether it's in depth experience with a particular file format, getting to grips with emulation, or writing a business case to fund your preservation work, it's unrealistic for a single person to have all the experience and answers needed. Getting support and advice from a subject specialist or simply someone who's been there, not succeeded first time, tried again, and learned the lessons can often make a real difference. That's where the DPC comes in. This is our guide to seeking help with your digital preservation challenges.



## DPC Consultancy

The DPC is now offering consultancy services for a trial period. This is a paid for service and is available at our discretion to all - members and non-members alike. Potential topics of work will be led by the DPC's expertise, including: maturity modelling, policy and strategy development, forward planning and bespoke digital preservation advice on a range of topics. We will seek to develop this offering in line with our Values, including: care for our members, resources, people and environment as well as maintaining our neutrality in respect to solutions, approaches, sectors, suppliers and vendors. To find out more and discuss your needs with us, please [visit our DPC Consultancy page](#) and/or contact: [karyn.williamson@dpconline.org](mailto:karyn.williamson@dpconline.org)

# Cross-sector Partnerships



Documentation

[About](#) [Blog](#) [Help](#) [Projects](#) [Developers](#)

[Help](#) [FAQ](#) [Docs](#) [Guides](#)

## GitHub and Software

[Enable a repository](#)

[Describe software](#)

[Archive software](#)

[Upload software manually](#)

[Archive a release from GitHub](#)

[Docs](#) / [GitHub and Software](#) / [Archive software](#) / [Archive a release from GitHub](#)

## Archive a release from GitHub

This guide will help you create a **Software** record to be archived in [Software Heritage](#). It assumes you have already enabled a repository.

If you need help to enable a repository, read the following guide:

- [Enable a repository](#)

~~1. precedents~~

2. challenges

~~3. opportunities~~

# What does it mean to preserve AI?

- Do we preserve AI in the artifactual sense - e.g., preserve records about how AI is developed?
- Do we adopt a bit-level preservation approach or do we need emulation as well?
- Are we focused on preserving models?
- Are we focused on preserving models *and* training data?
- Given copyright concerns, is it possible to preserve training data? E.g., Many Hugging Face repos do not provide sufficient access to training data to archive them, likely due to creators seeking to avert real or perceived legal risk.

# Who is mission-aligned to to preserve AI, long-term?

- In the United States, primarily research libraries operating in a higher education context.
  - Long term, multi-generational, often perpetual preservation commitments.
  - For better or worse, decentralized.
  - Commitments come with the understanding that preservation of many versions of AI helps individuals from many disciplines and sectors make sense of decision making and societal change over time.
- For-profit companies are not incentivized to invest in perpetual preservation of their products.
  - Taking a long view, they wax and wane according to the market.
  - With ~30 years of digital preservation experience, we have numerous use cases of companies that have come and gone whose decline often triggers rapid preservation response efforts from the LAM community.

# Who has the financial resources to preserve AI?

- In the United States higher education research libraries are a natural candidate but recent years weaken that position. A rapid weakening.
  - Weakening student enrollment and enrollment revenue
    - Demographic cliff, coming due
    - Current administration discourages international student enrollment
  - Federal funding uncertainty deeply impacting university research revenue
    - Eliminated funding agencies
    - Underperforming agencies due to layoffs and uncertainty
    - Drops in federal funding = drastic drops in library portion of F&A funds
  - Federal and private funders have largely turned toward nationalistic > international effort
  - Hiring freezes, and tough decisions about ongoing programmatic investment is negatively impacting preservation effort.

# Who has the financial resources to preserve AI?



[About](#) ▾

[Membership](#) ▾

[Groups](#) ▾

## The Future of NDSA

Dec 18, 2025

As NDSA continues to evolve, the Coordinating Committee and broader leadership team would like to close the year by sharing our vision for the future of the organization and where NDSA is heading in 2026. What follows is an edited version of the presentation given by NDSA Chair at the 2025 session “**The State of the NDSA.**”



[ABOUT SPN](#)

[EAASI RESEARCH ALLIANCE](#)

[JOIN SPN](#)

[WORKING GROUPS](#)

# Frequently Asked Questions: Software Preservation Network Sunset

~~1. precedents~~

~~2. challenges~~

3. opportunities

# Cross-Sector Partnerships



Documentation

[About](#)

[Blog](#)

[Help](#)

[Projects](#)

[Developers](#)

[Help](#)

[FAQ](#)

[Docs](#)

[Guides](#)

## GitHub and Software

[Enable a repository](#)

[Describe software](#)

[Archive software](#)

[Upload software manually](#)

[Archive a release from GitHub](#)

[Docs](#) / [GitHub and Software](#) / [Archive software](#) / [Archive a release from GitHub](#)

## Archive a release from GitHub

This guide will help you create a **Software** record to be archived in [Software Heritage](#). It assumes you have already enabled a repository.

If you need help to enable a repository, read the following guide:

- [Enable a repository](#)





# Cross-Sector Partnerships



**Hugging Face**

# Cross-Sector Partnerships

KB } nationale  
bibliotheek

 Catalogus  Zoeken  Werken bij  Contact **NL** EN

Onderzoeken & vinden ∨ Bezoek & lidmaatschap ∨ Ontdekken & bewonderen ∨ Over ons ∨

Home > GRIN (Global Resilient Information Network)

## GRIN (Global Resilient Information Network)

# Cross-Sector Partnerships

- GRIN Participants
  - Multiple national libraries
  - Academic research libraries
  - Nonprofit independent research libraries
  - Nonprofit scholarly technology providers
  - And growing ...

# Philanthropic and Government Funding

- Evolve digital preservation practice relative to AI preservation use cases
- Provide runway for nonprofit organizations to develop sustainability models for ongoing AI preservation practice
- Foster multi-sector collaborations that support AI preservation
- Foster multidisciplinary study of archived AI and its impact on society
- Foster international collaboration, counter to current nationalist trends

We have work to do aligning research libraries, AI service providers, philanthropy and policymakers so they can collaborate on long-term AI preservation.

Alignment will be key to resourcing a stable transition from maximally useful AI to an artifactual state available for future study.

AI can and must be preserved.



**Thank you.**

Thomas Padilla  
Associate Dean, Research and Learning  
University of Nebraska Lincoln

6.5.26