

# Digital preservation and archiving in the NFDI

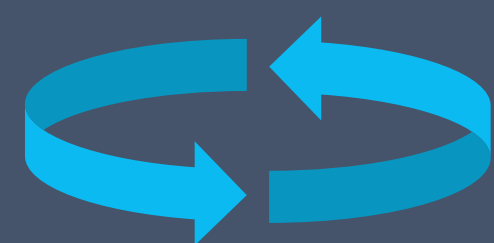
Between established standards, varying levels of awareness and approaches to open questions

## Why digital preservation matters

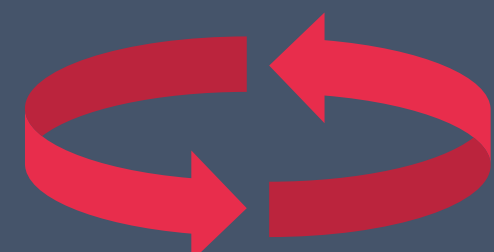
- Ensures **long-term reusability, accessibility, and reproducibility** of research data
- Prevents data loss which can compromise:
  - Verification of results
  - Citation tracking & provenance
  - Reuse in future research
- In order to ensure sustainable usability of research data, a digital preservation strategy appropriate to the respective research discipline is essential. (Wissenschaftsrat 2025 [8])

## What is digital preservation

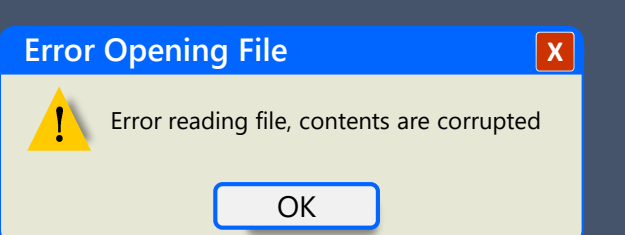
- “The act of **maintaining information**, Independently Understandable by a Designated Community, and with evidence supporting its Authenticity, over the Long Term” [1]
- Following **standards and best practices**: ISO 14721, CoreTrustSeal certification, nestor seal certification, ...
- Entails bitstream and content preservation [2]



**Bitstream preservation:** preservation of the bitstream of digital objects, e. g. storage monitoring, redundant storage, backups, checksums, ...



**Content preservation:** preservation of content of digital objects beyond changes in technology, standards, etc. E. g. file format migration, emulation, ...

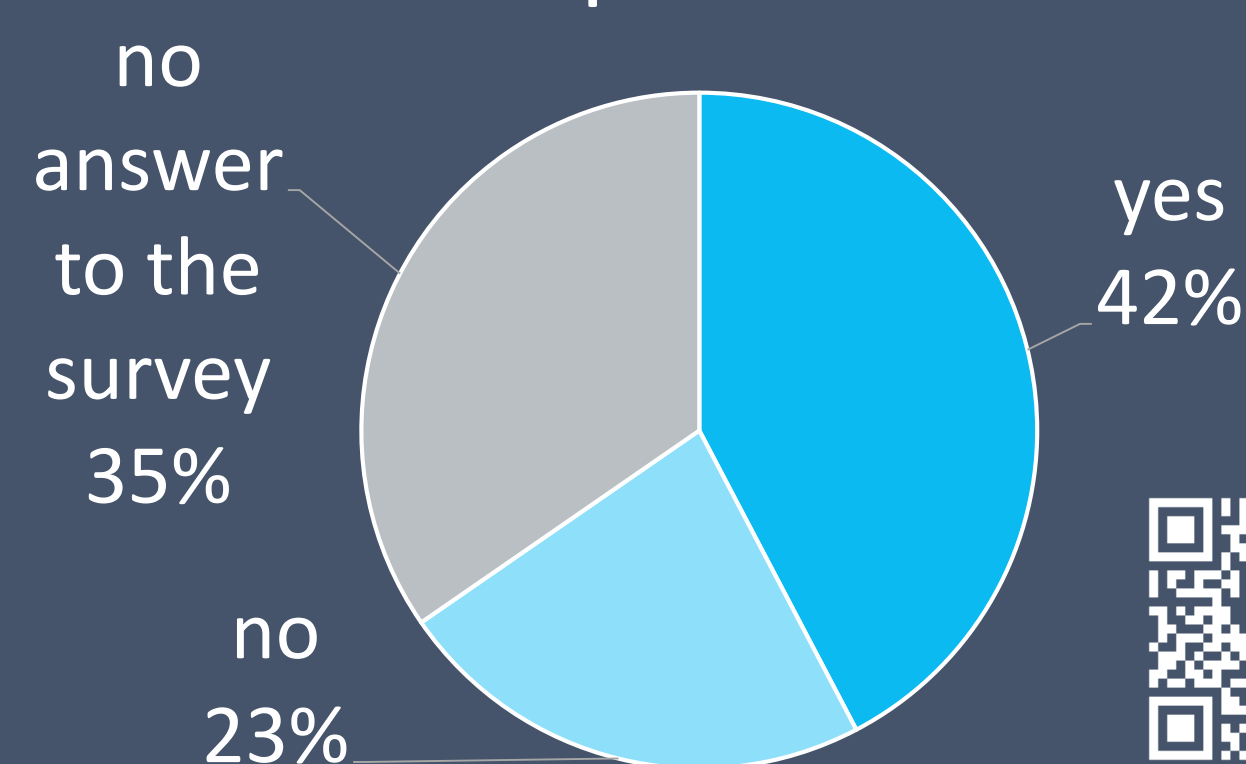


## Current landscape

### Challenges and survey findings



#### NFDI consortia offering bitstream preservation



#### NFDI consortia offering content preservation

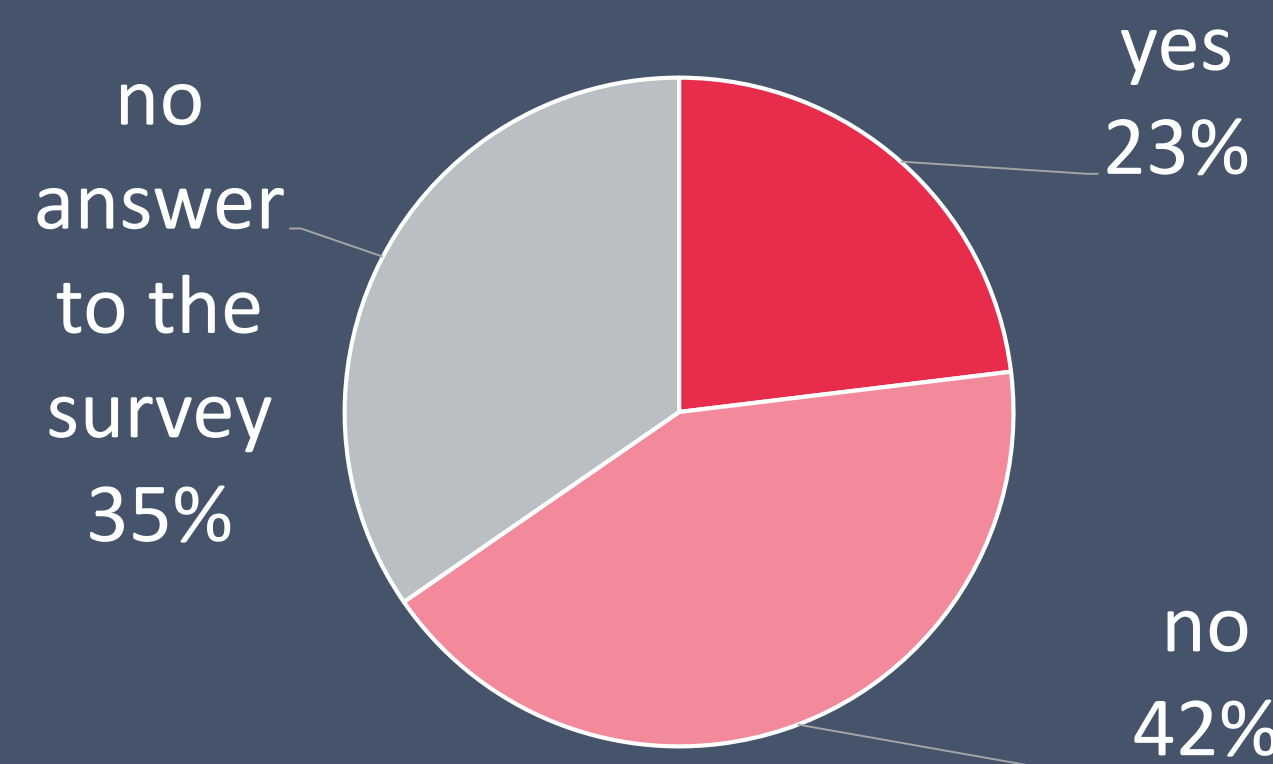


Fig. NFDI consortia offering bitstream and content preservation. Survey was conducted 2023 by WG LTA among consortia; 17 consortia of 26 responded. In case of a positive answer for content preservation, bitstream preservation was assumed as content preservation builds on it [3].

- 11 of 17 consortia expressed the need for a **digital preservation service** (incl. for specific data types) (survey 2023 [3])
- Criteria for **selecting data** for preservation (appraisal) needed [4]
- Clarification of long-term **responsibility** (public policies) needed [3, 4]
- Software preservation:** improvement of long-term reusability as an open question [5]
- Professionalizing data curation regarding **file errors** (e. g. format validation errors, format identification errors, insufficiently specified formats)

## NFDI working group LTA

- Long-Term Access and Preservation (LTA)
- Working group charter [7]
- In contact with nestor, EOSC



### Current projects

- Central information resource for **preservation status of published research data** (DataCite, digital preservation marker) [6]
- Integrating preservation in **DMP templates** (with DMP4NFDI)
  - Which research data will be preserved long-term how by whom?
  - Which data components should be preserved in the long-term?
  - Handling data along the life cycle: are rights, context information, dependencies, etc. sufficiently documented?
  - DMP templates reusable for all consortia



## Outlook and plans



### Digital preservation and archiving **consultation service** (funding pending)

- Clarify preservation needs further
- Consult NFDI consortia and existing research data services

### Idea exchange

- Additional **workshops** about digital preservation topics
- Clarify role of certified and trustworthy digital archives

### NFDI development

- Recommendations regarding standards and technical **Research Data Commons (RDC)** development

#### References

- [1] ISO 14721. See also OAIS Reference Model <https://public.ccsds.org/Pubs/650x0m3.pdf>
- [2] Object Levels of Preservation, Thibodeau, K. 2002. The State of Digital Preservation: An International Perspective <http://www.clir.org/wp-content/uploads/sites/6/pub107.pdf>
- [3] [10.5281/zenodo.11109480](https://doi.org/10.5281/zenodo.11109480)
- [4] E. g. digital preservation policy, see guidelines for the creation of a digital preservation policy <https://nbn-resolving.de/urn:nbn:de:0008-2014111006>
- [5] [10.5281/zenodo.15106332](https://doi.org/10.5281/zenodo.15106332)

#### References

- [6] <https://github.com/datacite/datacite-suggestions/discussions/149>
- [7] [10.5281/zenodo.6451455](https://doi.org/10.5281/zenodo.6451455)
- [8] <https://doi.org/10.57674/wcdc-6d36>

#### Contact [section-infra-wg-lta@lists.nfdi.de](mailto:section-infra-wg-lta@lists.nfdi.de)

#### Working group LTA organization team:

- Thomas Stäcker (TU-Darmstadt) [thomas.staecker@tu-darmstadt.de](mailto:thomas.staecker@tu-darmstadt.de)
- Svenia Pohlkamp (DNB) [S.Pohlkamp@dnb.de](mailto:S.Pohlkamp@dnb.de)
- Tobias Steinke (DNB) [T.Steinke@dnb.de](mailto:T.Steinke@dnb.de)
- Katharina Markus (ZB MED) [markus@zbmed.de](mailto:markus@zbmed.de)

Poster created by: Katharina Markus, with input by WG LTA