

# Modelling curation and preservation levels for trustworthy digital repositories

**PV2023, Geneva, Switzerland**  
**May 4, 2023**

Dr. Jonas Recker<sup>1</sup> | Hervé L'Hours<sup>2</sup> | Mari Kleemola<sup>3</sup>

<sup>1</sup>GESIS - Leibniz Institute for the Social Sciences, CoreTrustSeal Board

<sup>2</sup>UK Data Service, UK Data Archive, University of Essex, CoreTrustSeal Board

<sup>3</sup>Finnish Social Science Data Archive, Tampere University, CoreTrustSeal Board



This work is licensed under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

# Curation and Preservation

- Organizational framework (policies, resources, mission, etc.)
- Collection characteristics (e.g. volume, object types and formats, etc.)
- Designated Community needs, preferences, knowledge base



When it comes to “adequate” curation and preservation, one size does not fit all

→ depends on objects, community, objective, and more

# Examples of existing definitions and models

## Institution-specific

- [PANGAEA Levels of Curation](#)
- [MIT Levels of Preservation Commitment](#)
- [ICPSR Levels of Curation](#)



## Community-based

- [NDSA Levels of Preservation](#)
- [The DCN CURATE\(D\) Steps](#)
- CoreTrustSeal Levels of Curation
- [GFBio Curation Levels](#)

# Background

- Community agreement on definition of “levels of care” provided by a repository provide an important reference point, e.g. for
  - planning and conception of services,
  - transparent communication with stakeholders,
  - benchmarking and evaluation.

# CoreTrustSeal Levels of Curation 2019-2025

## From CoreTrustSeal Guidance

Level	Description
A.	Content distributed as deposited.
B.	Basic curation – e.g. brief checking, addition of basic metadata or documentation
C.	Enhanced curation – e.g. conversion to new formats during ingest, enhancement of documentation and metadata
D.	Data-level curation – as in C above, but with additional editing of deposited data

2019-2021, 2020-2022

- Curation understood as “adding value by enhancing content”
- Initial deposits unchanged, edits only on copies

2023-2025

- Curation as prerequisite for “assuring long-term accessibility and understandability of data as the needs of the Designated Community change”
- Initial deposits unchanged, edits only on copies
- Sufficient metadata for independent use by Designated Community
- Measures for active preservation in place

CoreTrustSeal Standards and Certification Board. (2019). CoreTrustSeal Trustworthy Data Repositories Requirements 2020–2022 (v02.00-2020-2022). Zenodo. <https://doi.org/10.5281/zenodo.3638211>

CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Requirements 2023-2025 (V01.00). Zenodo. <https://doi.org/10.5281/zenodo.7051012>



# Revising the CoreTrustSeal Levels of Curation (1)

- Integrate community consensus into a “core” level for trusted digital repositories
- Community consultation & requests for feedback
  - during two revisions of the CoreTrustSeal Requirements (2019, 2021)
  - CoreTrustSeal Discussion Paper on “Curation and Preservation Levels”<sup>1</sup>



# Revising the CoreTrustSeal Levels of Curation (2)

- Comments / suggestion received:
  - clarification of concepts and terminology
  - relationship between levels of curation and (expectations for) digital preservation measures not sufficiently defined
  - seemingly prescriptive of particular preservation strategies (normalization and migration)
  - strongly focused on curation at ingest

See, for example,

<https://www.coretrustseal.org/why-certification/review-of-requirements/>; Lindlar, Micky, & Rudnik, Pia. (2019). "Eye On CoreTrustSeal - Recommendations for Criterion R0 from Digital Preservation and Research Data Management Perspectives." iPRES2019.

<https://doi.org/10.5281/zenodo.3529423>; CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Revision Working Group Change Log and Associated Materials (v01.00). <https://doi.org/10.5281/zenodo.7051237>



# Revised curation & preservation levels

A. Conceptual preservation for understanding and reuse	In addition to B., the repository takes long-term responsibility that the <a href="#">data content and metadata can be independently understood</a> by the designated community.
B. Logical-Technical Curation	In addition to D. and/or C. the repository takes long-term responsibility for ensuring that the <a href="#">data and metadata can be rendered as required</a> by the designated community.
C. Initial Curation	In addition to Level D., if these criteria are not met the digital objects are <a href="#">curated by the repository</a> to meet the defined criteria..
D. Deposit Compliance	Data content and supporting metadata deposited are <a href="#">checked at the point of deposit for compliance with defined criteria</a> .
Z. Content distributed as deposited.	Data content and supporting metadata are <a href="#">distributed to users exactly as they are provided by depositors</a> . Data content and supporting metadata are stored for a given time period, or indefinitely.





# Next Steps

- “Curation & Preservation Levels” Discussion Paper
  - Current status is open consultation on [Version 1](#)
  - Next: Version 2 to reflect feedback from members of the LTDP TF
- View to present at RDA Maintenance Group for RDA as candidate for future CoreTrustSeal revisions
- Parallel discussions, including input to the LTDP TF and the FAIR IMPACT project about implications for object level metadata
  - Aligned metadata between stored/curated/preserved and the repository offering those services could provide valuable input into the emerging graphs of research metadata and make the current repository service offerings and digital object care levels more transparent.
- Seeking other areas where the levels could receive feedback or be candidates for adoption.



# Questions & Comments?

Contact

[jonas.recker@geis.org](mailto:jonas.recker@geis.org)

