# Revision of the Preservation Metadata DPC Report - A Preview

Bertrand Caron[1] Micky Lindlar[1]

[1]Technische Informationsbibliothek (TIB)
Langzeit Archivierung

2026-04-24

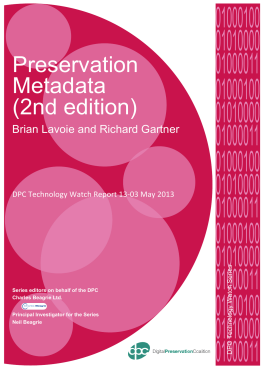
TIB LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

# What I'll cover today...

1. A Bit of Context: 2013-2026
2. My Proposal for the 3rd Edition

# The Previous Edition

Gartner, Richard, and Brian Lavoie. *Preservation Metadata (2nd Edition)*. Digital Preservation Coalition, 2013. https://doi.org/10.7207/twr13-03.



# What happened since 2013?

(Non-comprehensive list...)

- 2015: PREMIS 3.0

# What happened since 2013?

(Non-comprehensive list...)

- 2015: PREMIS 3.0
- 2016: Dappert, Angela, Rebecca Squire Guenther, and Sébastien Peyrard, eds. *Digital Preservation Metadata for Practitioners*. Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-43763-7.

# What happened since 2013?

(Non-comprehensive list...)

- 2015: PREMIS 3.0
- 2016: Dappert, Angela, Rebecca Squire Guenther, and Sébastien Peyrard, eds. *Digital Preservation Metadata for Practitioners*. Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-43763-7.
- 2017: Preservation Action Registry (PAR)

# What happened since 2013?

(Non-comprehensive list...)

- 2015: PREMIS 3.0
- 2016: Dappert, Angela, Rebecca Squire Guenther, and Sébastien Peyrard, eds. *Digital Preservation Metadata for Practitioners*. Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-43763-7.
- 2017: Preservation Action Registry (PAR)
- 2017-2018(?): Wikidata for Digital Preservation (WikiDP)

# What happened since 2013?

(Non-comprehensive list...)

- 2015: PREMIS 3.0
- 2016: Dappert, Angela, Rebecca Squire Guenther, and Sébastien Peyrard, eds. *Digital Preservation Metadata for Practitioners*. Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-43763-7.
- 2017: Preservation Action Registry (PAR)
- 2017-2018(?): Wikidata for Digital Preservation (WikiDP)
- 2018: PREMIS v. 3 ontology

# What happened since 2013?

(Non-comprehensive list...)

- 2015: PREMIS 3.0
- 2016: Dappert, Angela, Rebecca Squire Guenther, and Sébastien Peyrard, eds. *Digital Preservation Metadata for Practitioners*. Springer International Publishing, 2016. `https://doi.org/10.1007/978-3-319-43763-7`.
- 2017: Preservation Action Registry (PAR)
- 2017-2018(?): Wikidata for Digital Preservation (WikiDP)
- 2018: PREMIS v. 3 ontology
- 2021: Community-Owned Digital Preservation Tool Registry (COPTR)

TIB LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

# What happened since 2013?

(Non-comprehensive list...)

- 2015: PREMIS 3.0
- 2016: Dappert, Angela, Rebecca Squire Guenther, and Sébastien Peyrard, eds. *Digital Preservation Metadata for Practitioners*. Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-43763-7.
- 2017: Preservation Action Registry (PAR)
- 2017-2018(?): Wikidata for Digital Preservation (WikiDP)
- 2018: PREMIS v. 3 ontology
- 2021: Community-Owned Digital Preservation Tool Registry (COPTR)
- 2025: METS 2.0

TIB LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

What is in the 2nd Edition

- OAIS
- PREMIS Data Dictionary
- PREMIS Revisions
- PREMIS Outreach
- PREMIS and METS
- PREMIS Tools
- PREMIS Implementations

What is not in the 2nd edition

- Registries
- Identifiers
- RDF-based expressions
- Other standards
- Practical examples
- Preservation metadata outside of the scope of PREMIS

# What I'll cover today...

1 A Bit of Context: 2013-2026

2 My Proposal for the 3rd Edition

# What is Digital Preservation, After All?

The initial definition of "preservation metadata" by OCLC was "information that supports and documents the digital preservation process".

But what is "digital preservation"?

DPC defines it as "the series of managed activities necessary to ensure continued access to digital materials for as long as necessary".

Obviously, that contradicts PREMIS which "is not concerned with discovery and access" (Caplan, Priscilla and PREMIS Editorial Committee. 'Understanding PREMIS').

TIB LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

# What is Preservation Metadata, After All?

Depending on how we define the notion, we might cover a variety of subjects...

- Internal metadata?
- Rich descriptive metadata provided by the producer?
- Fixity information in manifests?
- Structured output of analysis tools?

My proposal for the scope of the DPC report:

**Definition**

Information in a structured, textual form

# What is Preservation Metadata, After All?

Depending on how we define the notion, we might cover a variety of subjects...

- Internal metadata?
- Rich descriptive metadata provided by the producer?
- Fixity information in manifests?
- Structured output of analysis tools?

My proposal for the scope of the DPC report:

**Definition**

Information in a structured, textual form, used by the Archive

MATION CENTRE
ND TECHNOLOGY
BRARY

# What is Preservation Metadata, After All?

Depending on how we define the notion, we might cover a variety of subjects...

- Internal metadata?
- Rich descriptive metadata provided by the producer?
- Fixity information in manifests?
- Structured output of analysis tools?

My proposal for the scope of the DPC report:

**Definition**

Information in a structured, textual form, used by the Archive for the physical and logical preservation of the digital object

# What is Preservation Metadata, After All?

Depending on how we define the notion, we might cover a variety of subjects...

- Internal metadata?
- Rich descriptive metadata provided by the producer?
- Fixity information in manifests?
- Structured output of analysis tools?

My proposal for the scope of the DPC report:

**Definition**

Information in a structured, textual form, used by the Archive for the physical and logical preservation of the digital object, and for their access over the longest possible period

# What is Preservation Metadata, After All?

Depending on how we define the notion, we might cover a variety of subjects...

- Internal metadata?
- Rich descriptive metadata provided by the producer?
- Fixity information in manifests?
- Structured output of analysis tools?

My proposal for the scope of the DPC report:

**Definition**

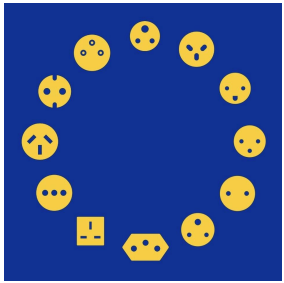
Information in a structured, textual form, used by the Archive for the physical and logical preservation of the digital object, and for their access over the longest possible period, preserved by the Archive along with the object itself.

MATION CENTRE
ND TECHNOLOGY
BRARY

# Standards in Scope

*"Standards are like toothbrushes. Everybody wants one but nobody wants to use anybody else's."*

- METS - version 2.0
- PREMIS - including examples in RDF!
- Portland Common Data Model (PCDM)
- Packaging formats (BagIt, OCFL) for fixity information
- Technical metadata standards
- RDF-based standards (OAI-ORE for structural information, PROV for provenance information, ODRL for rights information)
- and perhaps a few more...



TIB LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

# General Structure of the Document 1/2

- Intro - Differences between 2nd and 3rd editions
- Glossary
- Defining Users and User Scenarios
- Recording Information as Metadata
  - Recording Structural Information
  - Recording Provenance Information
  - Recording "Technical" Information
    - Recording Fixity Information
    - Recording Format Information
    - Recording Dependencies / Environment Information
    - Recording Quality and Risk Information
  - Recording Descriptive Information
  - Recording Source Information
  - Recording Rights Information

TIB LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

# General Structure of the Document 2/2

- Creating the Logical Model
  - Defining the Model Entities
  - Focus: Assigning Identifiers
  - Defining the Properties and Relationships of the Entities
- Selecting Metadata Solutions
  - Criteria for Selecting Metadata Schemas
  - Combining Existing Metadata Schemas
  - Using External Registries
  - Defining Application Profiles
- Serialising and Handling Metadata
  - Choosing Between Serialisation Options
  - Controlling Metadata
  - Mapping Metadata
  - Packaging Metadata
  - Querying Metadata

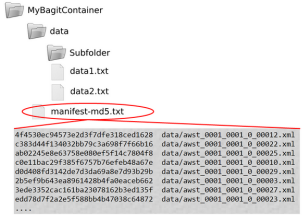
# Structure of the 'Recording XXX Information' Sections

- What? (definition)
- Why? (rationale)
- How (to produce)? (tools, sources)
- When (to produce)? (before transfer, at appraisal, at ingestion, after ingestion...)
- How (to encode)? (examples of alternative metadata structures)
- Which level (should it qualify)? (IP, object, object component, file, bitstream, etc.)

# Metadata Snippets

## For Fixity Information

**If the fixity information is stored in the BagIt packaging format**, it will be located in a specific file named "manifest-[algorithm name].txt".




## For Provenance Information

The same example can be expressed according to the recommendations for expressing PREMIS in RDF:[6]

```
<http://nri.library.ca/0912-0002.tif> a premis:File .

<0912-0001Event> a evType:ing ;
        prov:startedATime "2017-11-14T13:23:30Z" ;
        prov:endedATime "2017-11-14T13:26:11Z" ;
        evRelAgRole:aut <NRI> ;
        evRelAgRole:imp <JaneDoe> ;
        prov:used <http://nri.library.ca/0912-0001.tif> .

<JaneDoe> a prov:Person ;
        foaf:name "Doe, Jane" ;
        premis:note "Preservation specialist, Digital Initiatives Division" .

<NRI> a prov:Organization ;
        foaf:name "Not a Real Institution" .
```

The full example is available at https://github.com/lcnetdev/PREMIS/blob/master/examples/animal_antics.ttl.

# Questions? Suggestions?

- What would you like to learn from that document?
- Which are your favorite (learning) resources and papers on that topic?
- Is it too ambitious...?
- Do I need to include a presentation of the data model of each main standard before I present metadata snippets...
- ... or should I insert that between the "What/Why/When" and the "How/Which level"?

Publication due in July 2026!


TIB LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY