

Introducing openCost in DSpace

The path to a minimum viable implementation

Introduction

The earliest cost data in our repository Gutenberg Open Science (DSpace) date back to 2022. Currently, all data is stored internally, but growing interest in cost transparency has prompted a first try at implementing openCost.

openCost offers a standard format for cost data that can easily be distributed via OAI-PMH. Adopting a new metadata schema can be a daunting task. However, thanks to the University of Ulm's publicly shared XSLT (published at the DSpace Praxistreffen 2025), much of the groundwork was already laid for the implementation in DSpace.

Motivation

Instead of committing right away to a full realisation of the openCost schema, we focussed on the "low-hanging fruit":

1. add the data we already have in the repository.
2. Then, improve the metadata in the repository and
3. finally expand the openCost XML on the OAI interface.

How easily can our existing cost data be transformed to meet the minimum required fields of the openCost schema?

Outlook

- ✔ Get a minimal, but valid version of openCost up and running.
- Adjust current data to fit openCost standard (e.g., add ESAC IDs)
- Expand the DSpace metadata registry (e.g., actual invoice date)
- Add entities to model relationships between publications, invoices and contracts (e.g., cost types other than hybrid and gold OA)

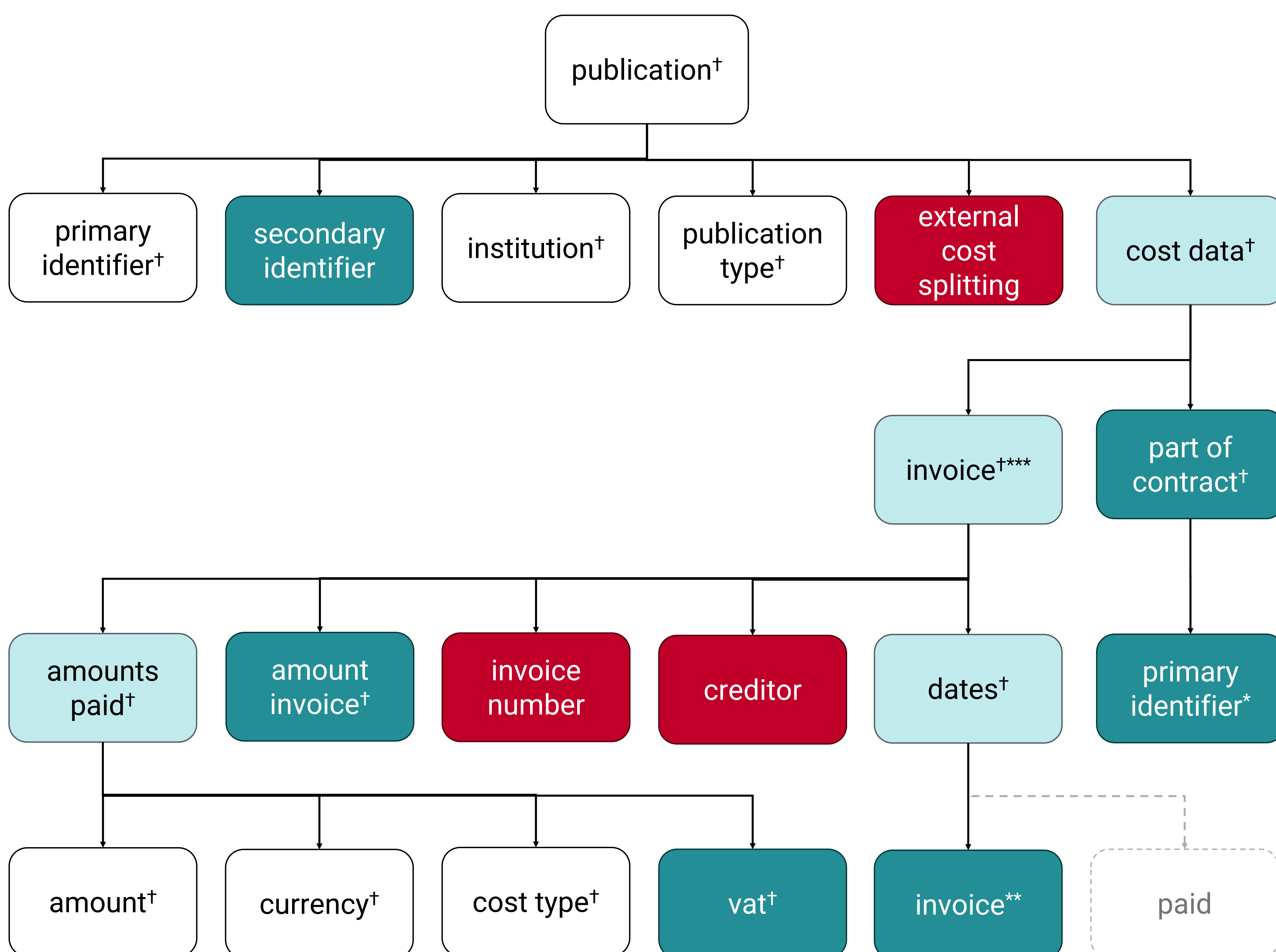
Step-by-step Approach

1. Assess the data in the repository.
2. Compare existing data against required elements.
3. Determine which elements could be filled with less detail.
4. Modify XSLT to use our DSpace metadata field names.
5. Compare existing data against optional elements.
6. Add new elements where data already exists in DSpace.
7. Re-evaluate the current state.
8. Plan new metadata fields to be added to DSpace.
9. Add the new metadata fields.
10. Repeat until openCost schema is fully implemented.

Conclusion

Using a minimal-requirements approach and the reusable XSLT from the University of Ulm, we produced a valid openCost OAI feed for the Gutenberg Open Science repository.

Current development in a test environment



† Required
* Currently only display names, not ESAC IDs
** Currently derived from DFG funding year, granularity: YYYY
*** Currently only one invoice relation per publication possible

○ Reused
● Added
● Added with placeholder data

Categorisations explained

○ Reused

The code could be reused with minimal changes. Some changes were necessary to adapt it to our data model. For example, changing the publication type mapping to the English DINI type labels.³

Wissenschaftlicher Artikel → Article

● Expanded

The code could be reused, but more data is available to be added. For example, the original currency of an invoice amount is already recorded in the repository.

● Added

The element did not exist previously and was newly added. For example, the secondary identifier did not exist yet. It was added for the types: doi, handle, urn, isbn and pmid.

● Added with placeholder data

No data for this element exists in the repository yet, however they will be added in the future. For example, the invoice number.

References

1. <https://openscience.ub.uni-mainz.de>
2. <https://cloudstore.uni-ulm.de/s/QaRpT5bP3r4mrSW>
3. <https://doi.org/10.18452/24148>