

Enhancing Cost Transparency: The Role of Persistent Identifiers in the openCost Metadata Schema



Authors: Bianca Schweighofer 0000-0002-9416-9311 (University Library of Regensburg)

Gernot Deinzer 0000-0002-7462-3847, Colin Sippl 0000-0002-8503-1740, Silke Weisheit 0000-0002-2609-4274 (University Library of Regensburg); Julia Bartlewski 0000-0001-5959-4999, Christoph Broschinski 0000-0003-1972-7587, Dirk Pieper 0000-0002-6083-9348 (Bielefeld University Library); Lisa-Marie Stein 0000-0001-7905-0462, Alexander Wagner 0000-0001-9846-5516 (DESY)

The openCost project

The aim of openCost is the creation of a technical infrastructure to comprehensively record all costs involved in scientific publishing and subsequently make them freely accessible by means of standardized interfaces.



This is supposed to enable cost transparency on an institutional, national, and international level.

The openCost GitHub repository:



- Contains the schema and its documentation, a tabular description of the schema, FAQs, an XSD file for formal verification, and a folder with valid sample files)
- Tracks the evolution of the schema in a transparent way
- Community participation via GitHub Issues

PIDs in the openCost metadata schema

PIDs improve the efficiency of data exchange processes and are therefore central to the openCost schema.

The use of established PIDs ensures clear referencing and interoperability of cost-related data, while promoting accessibility.



```
<xs:complexType name="publication_type">
  <xs:all>
    <xs:element name="primary_identifier" type="publication_primary_identifier" />
    <xs:element name="secondary_identifiers" type="publication_secondary_identifiers" minOccurs="0" />
    <xs:element name="institution" type="institution_type" />
    <xs:element name="publication_type" type="coar_publication_type" />
    <xs:element name="external_costsplitting" type="xs:boolean" minOccurs="0" />
    <xs:element name="cost_data" type="publication_cost_data_type" />
  </xs:all>
</xs:complexType>

<xs:complexType name="contract_type">
  <xs:all>
    <xs:element name="contract_name" type="non_empty_string" />
    <xs:element name="institution" type="institution_type" />
    <xs:element name="participation" type="participation_type" />
    <xs:element name="primary_identifier" type="contract_primary_identifier_type" />
    <xs:element name="secondary_identifiers" type="contract_secondary_identifiers_type" minOccurs="0" />
    <xs:element name="cost_data" type="contract_cost_data_type" />
  </xs:all>
</xs:complexType>
```

Further information:

openCost project: <https://opencost.de/en>

GitHub repository:

<https://github.com/opencost-de/opencost>

OpenAPC project: <https://openapc.net>

Electronic journals library (EZB):

<https://ezb.ur.de>



Project focus:

1 openCost Metadata Schema

Development of a standardized metadata schema to record, retrieve and map all publication costs of a scientific institution in a structured form. This also includes additional costs as well as processing fees or costs from transformative agreements. The schema is designed as a minimalist yet comprehensive data exchange format that ensures seamless integration with other formats.

2 Harvesting via OAI-PMH



Automatic exchange of this data via the established OAI-PMH interface. Via OAI-PMH, service providers (e.g., aggregators like OpenAPC and research funders) can harvest their publication outputs directly from the institutions.

3 Extension of the EZB



Extension of the Electronic Journals Library (EZB) to include special functions for displaying publication costs and their assumption in order to facilitate access to cost-relevant information. In addition, we want to provide a function that enables institutions to record institution-specific agreements such as memberships in the EZB administration.

- openCost: Schema for cost data, **DOI for bibliographic data**

- Automated subsequent use of standardized metadata using DOI via e.g. Crossref, DataCite

- **Problems:** Workflows for publication types without unique identifiers or publication types that are referenced via other identifiers

```
publication
  --- primary_identifier
  --- doi
  *either doi OR bibliographic_information
  --- bibliographic_information
    --- Title
    --- Publisher
    --- isPartOf
```

DOI



ESAC



- **ESAC:** central directory on **transformative agreements**

- **Collected data:** contracting parties, contract terms, transition mechanisms for fees, negotiated publication rights and workflow assessment; provision of **IDs**

- **Problems:** correct assignment; no machine-readable interface; no allocation between contract, covered journals and participating institutions; business models other than transformative agreements such as Diamond OA or SCOAP³ are not covered