Publication CERIF mapping

Definitions

```
header = normdoc / part[name='record'] // header
```
```
knaw_long = normdoc / part[name='record'] // metadata / normalized / knaw_long
```

Internal Identifier

```
header / identifier
```

Type

```
knaw_long // genre
```

If not provided, use 'text (http://purl.org/coar/resource_type/c_18cf)'

mapping:

knaw_long pub_types/genres	COAR url
annotation	annotation (http://purl.org/coar/resource_type/c_1162)
article	journal article (http://purl.org/coar/resource_type/c_6501)
bachelorthesis	bachelor thesis (http://purl.org/coar/resource_type/c_7a1f)
book	book (http://purl.org/coar/resource_type/c_2f33)
bookpart	book part (http://purl.org/coar/resource_type/c_3248)
bookreview	book review (http://purl.org/coar/resource_type/c_ba08)
conferencepaper	conference paper (http://purl.org/coar/resource_type/c_5794)
contributiontoperiodical	contribution to journal (http://purl.org/coar/resource_type/c_3e5a)
doctoralthesis	doctoral thesis (http://purl.org/coar/resource_type/c_db06)
researchproposal	research proposal (http://purl.org/coar/resource_type/c_baaf)

lecture	lecture (http://purl.org/coar/resource_type/c_8544)
masterthesis	master thesis (http://purl.org/coar/resource_type/c_bdcc)
patent	*should not occur for publications in CERIF*
preprint	preprint (http://purl.org/coar/resource_type/c_816b)
report	report (http://purl.org/coar/resource_type/c_93fc)
studentthesis	thesis (http://purl.org/coar/resource_type/c_46ec)
technicaldocumentation	technical documentation (http://purl.org/coar/resource_type/c_71bd)
workingpaper	working paper (http://purl.org/coar/resource_type/c_8042)
conferenceobject	conference object (http://purl.org/coar/resource_type/c_c94f)
other	text (http://purl.org/coar/resource_type/c_18cf)
conferenceitem	conference object (http://purl.org/coar/resource_type/c_c94f)
conferenceitemnotinproceedings	conference object (http://purl.org/coar/resource_type/c_c94f)
conferenceposter	conference poster (http://purl.org/coar/resource_type/c_6670)
conferenceproceedings	conference proceedings (http://purl.org/coar/resource_type/c_f744)
reportpart	report part (http://purl.org/coar/resource_type/c_ba1f)
review	review (http://purl.org/coar/resource_type/c_efa0)

Language

`knaw_long // language`

Title

`knaw_long // titleInfo / title`

Some `titleInfo` have a `xml:lang="en"` attribute, but this is always a duplicate of the title without `xml:lang`.
Use the `titleInfo` without `xml:lang` attribute.

Subtitle

`knaw_long // titleInfo / subtitle`

Some `titleInfo` have a `xml:lang="en"` attribute, but this is always a duplicate of the title without `xml:lang` .
Use the `titleInfo` without `xml:lang` attribute.

PublishedIn

Not possible, because we don't have the identifier to the 'published_in'

`knaw_long // relatedItem[type='host'] / titleInfo / title`

`type` attribute can have values 'host', 'series' and others, as listed on GitHub.
Not sure if we should only have use `type="host"` or if the others are allowed as well.

PartOf

Not recorded

PublicationDate

`knaw_long // dateIssued / parsed`

Number

Not recorded

Volume

`knaw_long // relatedItem[type='host'] / part / volume`

Issue

`knaw_long // relatedItem[type='host'] / part / issue`

Edition

Not recorded

StartPage

`knaw_long // relatedItem[type='host'] / part / start_page`

EndPage

`knaw_long // relatedItem[type='host'] / part / end_page`

DOI

`knaw_long // publication_identifier[@type='doi']`

Available identifiers:

- scopus
- pmid
- pure
- purl
- wos
- isbn
- nbn
- doi
- handle

Handle

`knaw_long // publication_identifier[@type='handle']`

PMCID

`knaw_long // publication_identifier[@type='pmid']`

Not sure if this is the same!

ISI-Number

`knaw_long // publication_identifier[@type='wos']`

ISI = Institute for Scientific Information, from Web of Science; might be this number

SCP-Number

`knaw_long // publication_identifier[@type='scopus']`

ISSN

`knaw_long // relatedItem / publication_identifier[@type='issn']`

medium

Not recorded

ISBN

`knaw_long // publication_identifier[@type='isbn']`

The ISBN values are formatted in some different ways.

- 978-0-19-957592-3
- 978-90-6171-998-4
- 9064648778
- 978 90 5931 068 1
- 9789057305368
- 9789460041204
- 9789057306877

Notice that CERIF specified a number of allowed regexes

medium

Not recorded

URL

`knaw_long // location_url`

URN

`knaw_long / persistentIdentifier`

Authors

Narcis doesn't record the PRS identifiers of the authors in `knaw_long` **. If a** `nameIdentifier` **is given in** `knaw_long` **, we could do a lookup in VSOI and add the PRS identifier. Only then can we add the authors to this CERIF format.**

Narcis aggregates and doesn't have a PRS identifier for each person that is listed as author.

Editors

Narcis doesn't record the PRS identifiers of the editors in `knaw_long` **. If a** `nameIdentifier` **is given in** `knaw_long` **, we could do a lookup in VSOI and add the PRS identifier. Only then can we add the authors to this CERIF format.**

Narcis aggregates and doesn't have a PRS identifier for each person that is listed as editor.

Publishers

No identifier available to link with.

License

`knaw_long // rightsDescription`

- *Not sure if this element can be added though. (1) I haven't seen it in use yet in the* `knaw` *store and (2) I'm not sure if we can offer both a* `value` *and a* `scheme` *for this.*
- *Can we map our values to the expected scheme? https://www.openaire.eu/cerif-profile/vocab/LicenseTypes#*

Subject

This requires a scheme for the subject, which we don't have available here or is not used. In the scheme of `knaw_long` **it says that a** `topicValue` **might contain a** `valueURI` **attribute, however, I can't find it in the KNAW store.**

Keyword

`knaw_long // subject / topic / topicValue`

(might be that part of these belong in Subject, but as long as they don't have a URL, we can't do anything with it)

Abstract

`knaw_long // abstract`

Status

Probably not recorded, but also there is no clear documentation or examples on what kind of values/URIs are expected here.

OriginatesFrom

This would require a link between these publications and the project(s) or funding(s) they originate from. Although some `grantAgreements` **are provided in** `knaw_long` **, I'm not clear on whether we have a 'referential transparent identifier' pointing to that funding.**

PresentedAt

Narcis doesn't record any `Event` **information that would be required for this entry.**

OutputFrom

Narcis doesn't record any `Event` **information that would be required for this entry.**

Coverage

Narcis doesn't record any `Event` **information that would be required for this entry.**

References

Not recorded

ns4:Access

`knaw_long / accessRights`

mapping:

knaw_long	CERIF
openAccess	open access (http://purl.org/coar/access_right/c_abf2)
embargoedAccess	embargoed access (http://purl.org/coar/access_right/c_f1cf)
restrictedAccess	restricted access (http://purl.org/coar/access_right/c_16ec)
closedAccess	metadata only access (http://purl.org/coar/access_right/c_14cb)