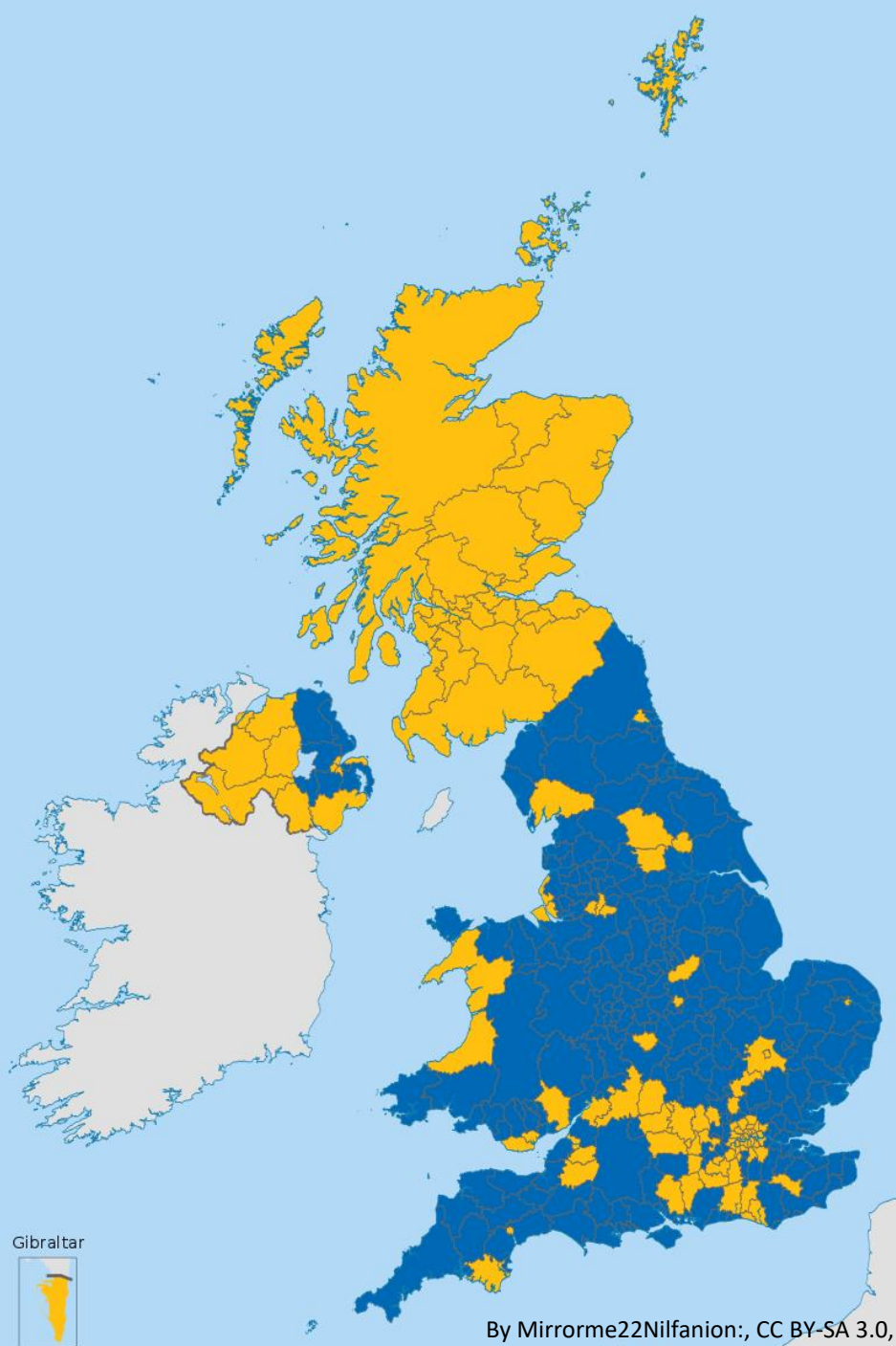


OpenAIRE CERIF-XML in Pure

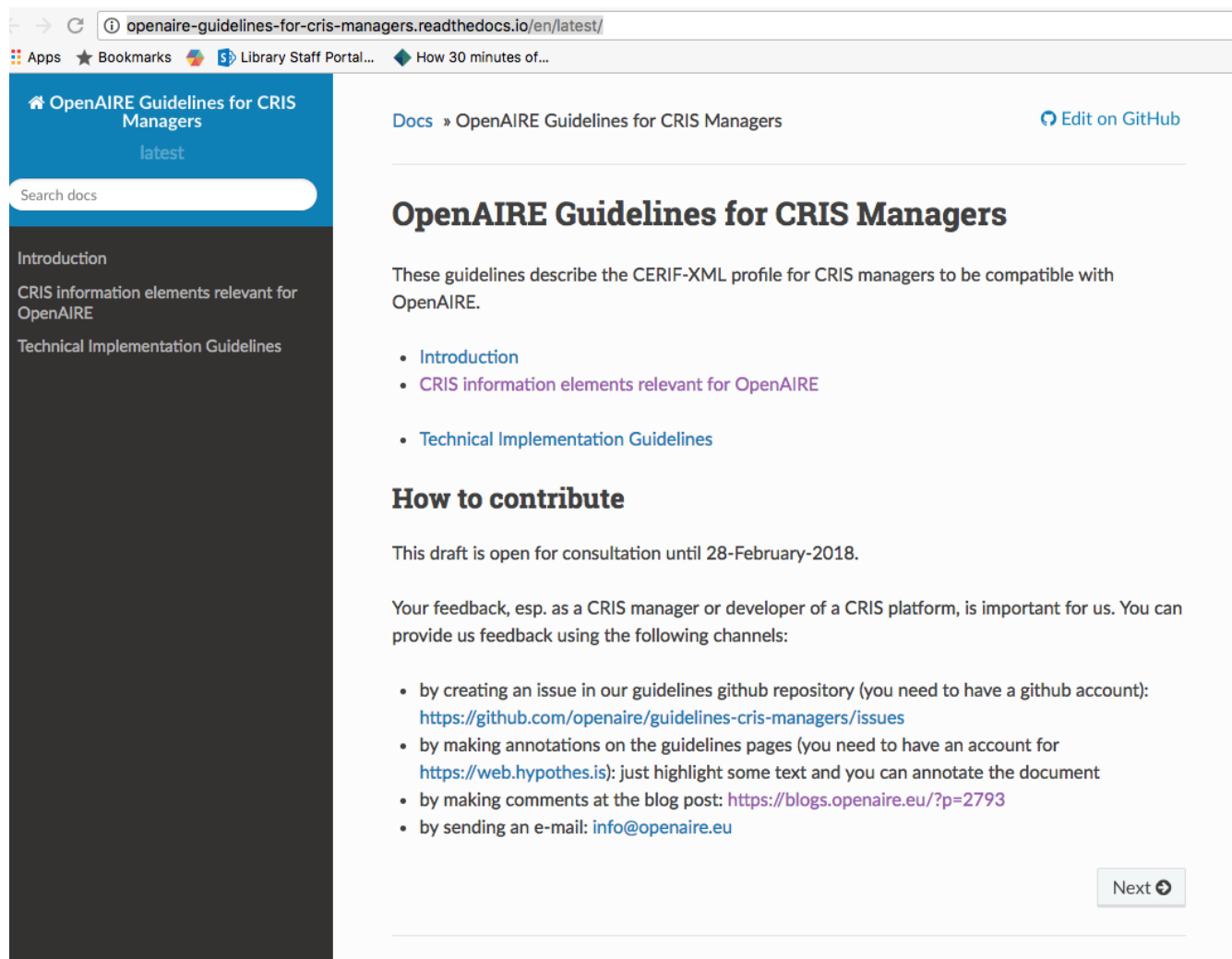
Anna Clements
euroCRIS & University of
St Andrews, UK
@AnnaKClements

ORCID: [0000-0003-2895-1310](https://orcid.org/0000-0003-2895-1310)

Jan Dvorak
euroCRIS & Czech
Technical University,
Prague

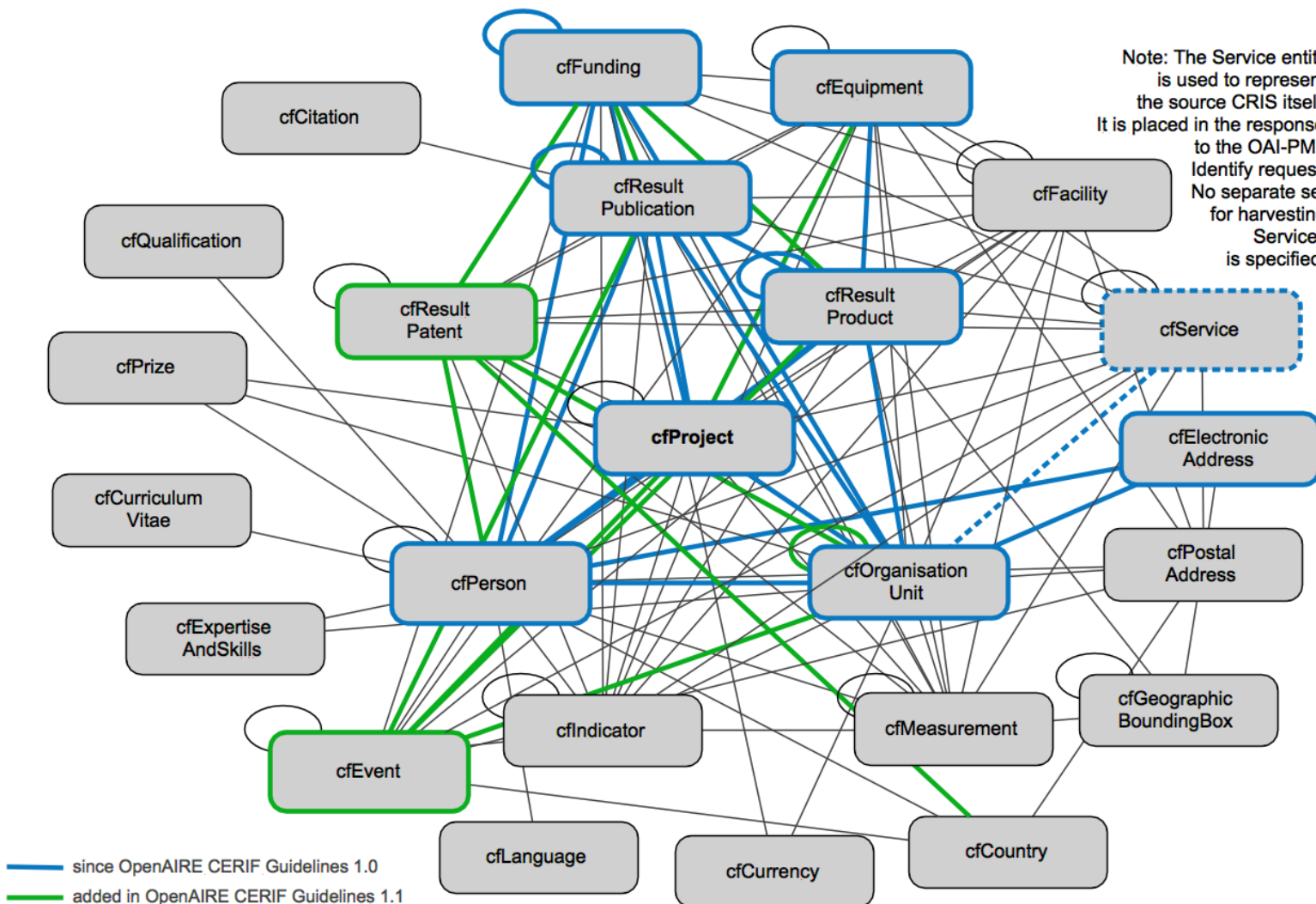


OpenAIRE CERIF-XML guidelines



The screenshot shows a web browser displaying the 'OpenAIRE Guidelines for CRIS Managers' page. The browser's address bar shows the URL 'openaire-guidelines-for-cris-managers.readthedocs.io/en/latest/'. The page has a blue header with the title 'OpenAIRE Guidelines for CRIS Managers' and a 'latest' version indicator. A search bar is present below the header. The left sidebar contains a navigation menu with 'Introduction', 'CRIS information elements relevant for OpenAIRE', and 'Technical Implementation Guidelines'. The main content area has a breadcrumb trail 'Docs » OpenAIRE Guidelines for CRIS Managers' and a link to 'Edit on GitHub'. The title 'OpenAIRE Guidelines for CRIS Managers' is prominently displayed. Below it, a paragraph states: 'These guidelines describe the CERIF-XML profile for CRIS managers to be compatible with OpenAIRE.' A bulleted list follows: '• Introduction', '• CRIS information elements relevant for OpenAIRE', and '• Technical Implementation Guidelines'. A section titled 'How to contribute' contains the text: 'This draft is open for consultation until 28-February-2018.' and 'Your feedback, esp. as a CRIS manager or developer of a CRIS platform, is important for us. You can provide us feedback using the following channels:'. A bulleted list provides feedback channels: '• by creating an issue in our guidelines github repository (you need to have a github account): <https://github.com/openaire/guidelines-cris-managers/issues>', '• by making annotations on the guidelines pages (you need to have an account for <https://web.hypothes.is>): just highlight some text and you can annotate the document', '• by making comments at the blog post: <https://blogs.openaire.eu/?p=2793>', and '• by sending an e-mail: info@openaire.eu'. A 'Next' button with a right arrow is at the bottom right of the content area.

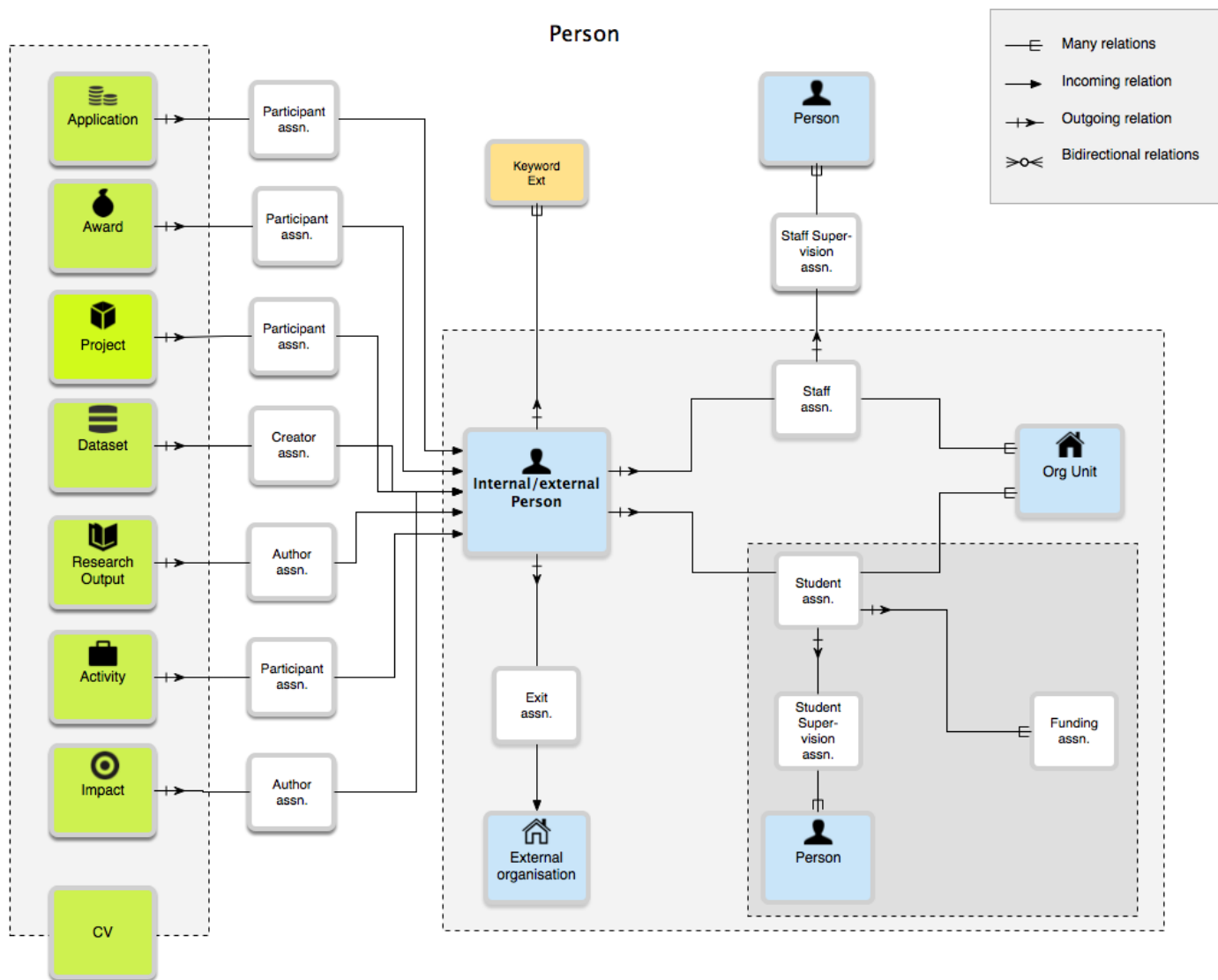
CERIF elements



Developing specification: mapping content types

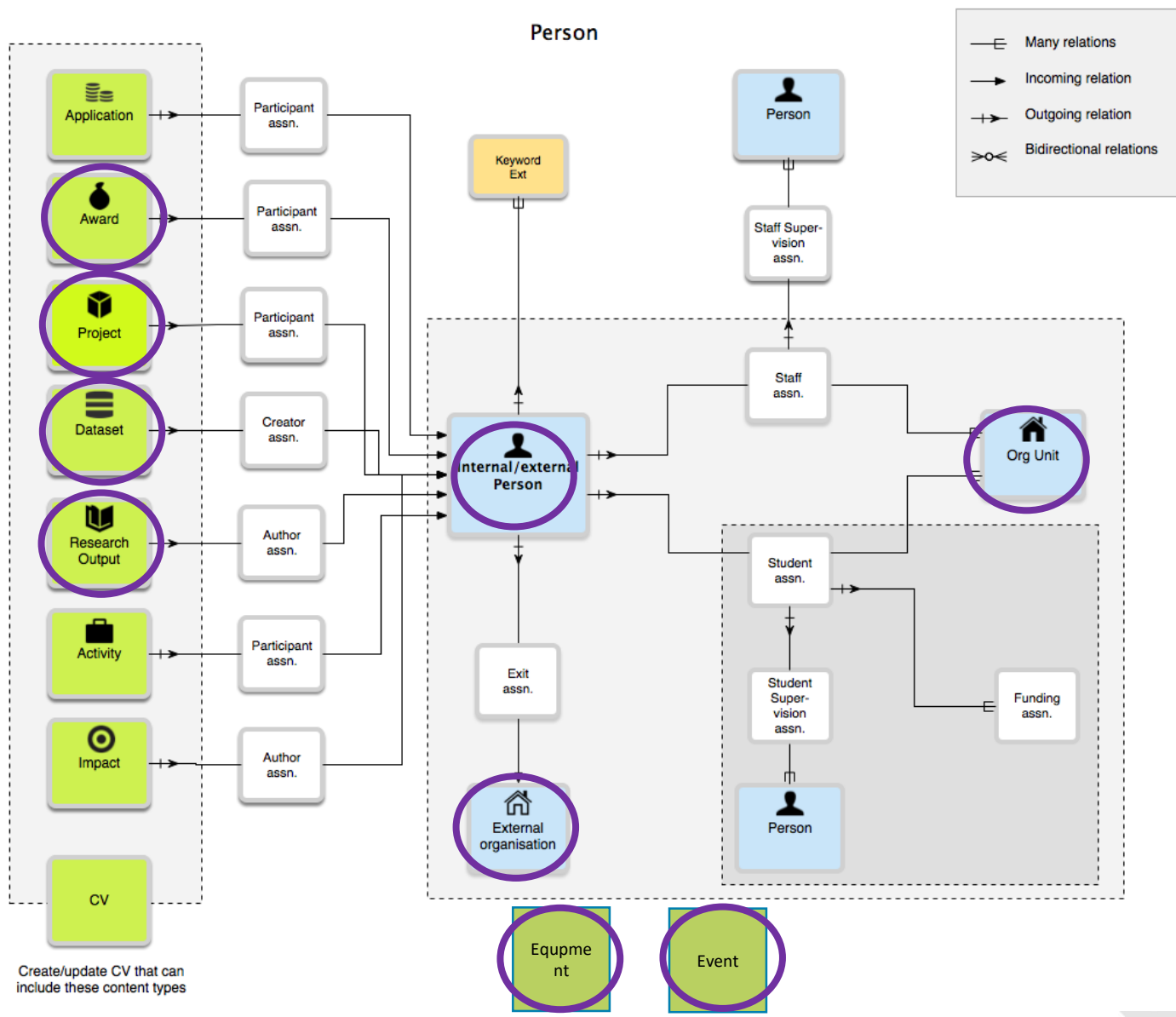
OpenAire CERF-XML
Publications
Products
Patent
Person
OrgUnit
Project
Funding
Service
Equipment
Event

Pure data model



Create/update CV that can include these content types

Pure data model



Developing specification: mapping content types

OpenAire CERF-XML	Pure 5.11 - suggested content types
Publications	Research Output
Products	Dataset
Patent	Template within Research Output
Person	Person, External person
OrgUnit	Organisational unit, External organisation, Publisher
Project	Project
Funding	Award
Service	??
Equipment	Equipment/facility
Event	Event

UK Pure User Group



Meet three times a year

How many members – 40+

Special interest groups:

- REF2021
- Repositories
- Research data
- Interoperability
- Theses

Short-life working groups:

- GDPR
- OpenAire

User-led

OpenAIRE Short Life Working Group

Aim : Provide mapping of Pure data model to OpenAire CERIF-XML guidelines and for CRIS Managers

Elsevier have committed to implementing these guidelines in PURE for the October 2018 (v 5.13). To meet this deadline, Elsevier have asked for input from euroCRIS and Pure Customers to help develop the specification mapping the data held in Pure to the new OpenAire CERIF-XML guidelines. A short life working group (SLWG) will be setup to deliver this specification.

SLWG Members

- Anna Clements (St Andrews, Chair)
 - Federica Fina (St Andrews)
 - Patricia Herterich (Birmingham)
 - Kara Jones (Bath)
 - Nykhola Strong (Aberdeen)
 - Anne Thorst Melbye (South Denmark)
 - Bill Worthington (Hertfordshire)
- Jan Dvorak (euroCRIS & Czech Technical University)
Alex Kujath (Elsevier)

Project timeline 2018

Feb	Pure dev version made available [5.11] with access to API, accompanying documentation and complete dataset for OpenAire mapping	Elsevier /Alex Kujath (AK)
Feb	Provide spreadsheet format of OpenAire CERIF-XML guidelines	euroCRIS/Jan Dvorak (JD)
Mar	Establish SLWG	euroCRIS/Anna Clements (AC)
Apr - Jun	Develop specification	SLWG
July	Implement specification	Elsevier
Aug	Feedback on implementation	SLWG
Sep	Test/Mature product	SLWG/Elsevier
Oct	Release 5.13.x	Elsevier

Developing specification: use of Pure web service

Pure API 511

This is the Pure Web Service. Listed below are all available endpoints, along with a short description.

In order to use the Pure Web Service, you must enter an API key. These are generated in the Administrator tab of Pure, and issued with a given set of available endpoints. To enter your API key and begin your use, press the 'Authorize' button to at the top of this page. You are then presented with two options for entering the API key: Use the API key in query format. Use the API key in a header. For further documentation, see [API Documentation](#).

A new version of the API is released with each major version of Pure, and remains available for one year. This version is no longer available in Pure 5.15. The old web service is deprecated, but still available [here](#), and it will no longer be available in Pure 5.13.

Created by Elsevier

activities	Show/Hide	List Operations	Expand Operations										
applications	Show/Hide	List Operations	Expand Operations										
author-collaborations	Show/Hide	List Operations	Expand Operations										
awards	Show/Hide	List Operations	Expand Operations										
changes	Show/Hide	List Operations	Expand Operations										
classification-schemes	Show/Hide	List Operations	Expand Operations										
concepts	Show/Hide	List Operations	Expand Operations										
curricula-vitae	Show/Hide	List Operations	Expand Operations										
datasets	Show/Hide	List Operations	Expand Operations										
GET /datasets Lists all datasets													
POST /datasets Complex operation for datasets													
GET /datasets/{id} Get dataset													
Implementation Notes Get dataset with specific ID (path parameter).													
Response Content Type application/xml													
Parameters <table><thead><tr><th>Parameter</th><th>Value</th><th>Description</th><th>Parameter Type</th><th>Data Type</th></tr></thead><tbody><tr><td>id</td><td>(required)</td><td>ID of the desired dataset</td><td>path</td><td>string</td></tr></tbody></table>				Parameter	Value	Description	Parameter Type	Data Type	id	(required)	ID of the desired dataset	path	string
Parameter	Value	Description	Parameter Type	Data Type									
id	(required)	ID of the desired dataset	path	string									

Developing specification: google sheets : orgunit draft



OpenAIRE Guidelines 1.1 mapping table - draft1



File Edit View Insert Format Data Tools Add-ons Help

Last edit was 5 minutes ago



	A	B	C	D	E
1	This is a draft table for mapping your CRIS onto the OpenAIRE Guidelines for CRIS Managers 1.1 CERIF profile				
2	The XML Schema for the OpenAIRE CERIF profile 1.1				
3	XPath (purely structural XML elements in <i>italic</i>)	Documentation	Cardinality	Populate from	Mapping Notes
16	OrgUnit/@id		exactly one	UUID	
17	OrgUnit/Acronym	The acronym of the organisation unit	optional	/organisational-units/organisationalUnit/nameVariants type="Short name"	
18	OrgUnit/Name	The name of the organisation unit	optional, possibly repeated if in different languages	/external-organisations/externalOrganisation/name /organisational-units/organisationalUnit/name	There is no postal address in the schema. In some external orgs in St Andrews Pure we have the full postal address in the name field. In others it is split out into separate address fields.
19	OrgUnit/Identifier	An identifier of the organisation unit	optional, possibly repeated	/external-organisations/externalOrganisation/uuid /external-organisations/pureId /organisational-units/organisationalUnit/addresses/address id	Should be Ringgold or FundRef ID any other shared identifier. Ask Alex what plans are to add orgids to Pure. Alex: In practical terms this looks like it will be the URL of the website of the organisations, see: https://github.com/openaire/guidelines-cris-managers/blob/master/samples/openaire_cerif_xml_example_orgunits.xml
20	OrgUnit/ElectronicAddress	An electronic address associated with the organisation unit	optional, possibly repeated	/external-organisations/externalOrganisation/email	For OrgUnit - email appears in UI but can't find in API
21	OrgUnit/PartOf	Chain up to the larger unit that encompasses this unit	optional	/organisational-units/organisationalUnit	Relationships can be seen for Organisational Units in the PURE UI, I couldn't find them exposed in the API. We also have to agree how much of the hierarchy should be included. For external organisations we don't have concept of hierarchy in Pure; often the full postal address is in the Name field. Sometimes this is split out into separate address fields .. but this is not a hierarchy - Update: PH found them now and added path accordingly

Developing specification: google sheets : orgunit *final*

2	The XML Schema for the OpenAIRE CERIF profile 1.1				
3	XPath (purely structural XML elements in <i>italic</i>)	Documentation	Cardinality	Skip	Populate from
17	Person/Affiliation/OrgUnit		exactly one		Mapping Notes
18	OrgUnit	Organisation Unit: an organisation, a unit therein, a committee or any other group of people that has a collective g			
19	OrgUnit/@id		exactly one		Internal and external organisations
20	OrgUnit/Acronym	The acronym of the organisation unit	optional		UUID of the record
21	OrgUnit/Type			1	Short name of internal organisation and acronym for external organisations
22	OrgUnit/Name	The name of the organisation unit	optional, possibly repeated if in differ		Name of the organisation
23	OrgUnit/Identifier	An identifier of the organisation unit	optional, possibly repeated		List all default IDs
24	OrgUnit/ElectronicAddress	An electronic address associated w	optional, possibly repeated		Webaddress of the internal organisation
25	OrgUnit/PartOf	Chain up to the larger unit that enc optional			
26	OrgUnit/PartOf/DisplayName		optional	1	Only populated for internal organisations
27	OrgUnit/PartOf/OrgUnit		exactly one		All direct parents organisation will be listed

Publication/date draft



OpenAIRE Guidelines 1.1 mapping table - draft1

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive



SHARE



	A	B	C	D	E
1	This is a draft table for mapping your CRIS onto the OpenAIRE Guidelines for CRIS Managers 1.1 CERIF profile				
2	The XML Schema for the OpenAIRE CERIF profile 1.1				
3	XPath (purely structural XML elements in <i>italic</i>)	Documentation	Cardinality	Populate from	Mapping Notes
106	Publication/PublicationDate	The date the publication appeared	optional	/contributiontype/publicationStatuses/publicationStatus/publicationDate/year /contributiontype/publicationStatuses/publicationStatus/publicationDate/month /contributiontype/publicationStatuses/publicationStatus/publicationDate/day	we might only reveal this element if /publicationStatus = published or e-pub ahead of print? this works for published journal articles but what about other cases? e.g. pre-prints, post-prints where pure or a local repository is the publication medium? should we look to the visibility/'available on date' of an open pre-publication manuscript in the way REF visibility is calculated? (ugh!) Agreed just rely on same visibility rules as for API
107	Publication/Number	The number of the publication (e.g. Article Number)	optional	/contributiontype/articleNumber	
108	Publication/Volume	The volume of the publishing channel where this publication appeared	optional	/contributiontype/volume	
109	Publication/Issue	The issue of the publishing channel where this publication appeared	optional	/contributiontype/journalNumber	
110	Publication/Edition	The edition of the publication	optional		
111	Publication/StartPage	The page where this publication starts, in case the publishing channel or containing publication has numbered pages	optional	/contributiontype/pages the first integer when the content takes the form integer[separator-nonintegers]integer eg 100-101, 100 - 101, 100:101 etc.	
112	Publication/EndPage	The page where this publication ends, in case the publishing channel or containing publication has numbered pages	optional	/contributiontype/pages the second integer when the content takes the form integer[separator-nonintegers]integer eg 100-101, 100 - 101, 100:101 etc.	
113	Publication/DOI		optional	/contributiontype/electronicVersions/electronicVersion	

Publication/date *final*

3	XPath (purely structural XML elements in <i>italic</i>)	Documentation	Cardinality	Skip	Populate from	Mapping Notes
116	Publication/PublicationDate	The date the publication appeared	optional		Will show published date, if available. If not, e-pub ahead of print date will be used. No other dates will be displayed	
117	Publication/Number	The number of the publication (e.g.	optional		Article number for contribution to journals	
118	Publication/Volume	The volume of the publishing channel	optional		Volume for contribution to journal and contribution to specialist types	contributionToJournal, contributionTo
119	Publication/Issue	The issue of the publishing channel	optional		Issue number for contribution to journal and contribution to specialist types	contributionToJournal, contributionTo
120	Publication/Edition	The edition of the publication	optional		Edition for book/report and contribution to book/anthology/report/conference proceeding types	bookAnthology, contributionToBookA
121	Publication/StartPage	The page where this publication starts	optional		Start page where available	
122	Publication/EndPage	The page where this publication ends	optional		End page where available	
123	Publication/DOI		optional		DOI	First DOI marked as "open" will be used
124	Publication/Handle		optional		Handle.net	all contribution types
125	Publication/PMCID		optional		"PubMedCentral" additional publication ID	all contribution types
126	Publication/ISI-Number		optional		"WOS" additional publication ID	all contribution types

Publication/type mapping p1

	A	B	C	D	E	F	G	H	I
1	Types of research output				Versions				
2	Base URI: /dk/atira/pure/researchoutput				dk.atira.pure.core:core:5.11.0-1		Updated with comments from George MacGregor 13/7/18		
3	Classifications:				dk.atira.pure.templates:templates:5.11.0-1				
4					dk.atira.pure.base-uk:base:5.11.0-1				
5					dk.atira.pure.base-uk.cust-uk-unifiedprojectdemo:cust:5.11.0-1				
6									
7	PURE 5.11.0-1				COAR Resource Type Vocabulary Version 2.0				
8	Type	SubType			Term	Publication or Product	Code/URL	Definition	
9	Contribution to journal								
10		Article		y	journal article	Publication	http://purl.org/coar/resource_type/c_6501	An article on a particular topic and published in a journal issue. (adapted from fabio)	
11		Meeting abstract		y	contribution to journal	Publication	http://purl.org/coar/resource_type/c_3e5a	A contribution to a journal denotes a work published in a journal. If applicable sub-terms sh	
12		Letter		y	letter to editor	Publication	http://purl.org/coar/resource_type/c_545b	A letter to the editor is a letter sent to a periodical about issues of concern from its readers. (	
13		Comment/debate		y	contribution to journal	Publication	http://purl.org/coar/resource_type/c_3e5a	A contribution to a journal denotes a work published in a journal. If applicable sub-terms sh	
14		Book/Film/Article review		y	review	Publication	http://purl.org/coar/resource_type/c_efa0	A review of others' published work.	
15		Literature review		y	contribution to journal	Publication	http://purl.org/coar/resource_type/c_3e5a	A contribution to a journal denotes a work published in a journal. If applicable sub-terms sh	
16		Editorial		y	editorial	Publication	http://purl.org/coar/resource_type/c_b239	A brief essay expressing the opinion or position of the chief editor(s) of a (academic) journa	
17		Special issue		y	contribution to journal	Publication	http://purl.org/coar/resource_type/c_3e5a	A contribution to a journal denotes a work published in a journal. If applicable sub-terms sh	
18		Review article		y	review article	Publication	http://purl.org/coar/resource_type/c_dcae04bc	A review article is a secondary source, that is, it is written about other articles, and does no	
19		Short survey		y	contribution to journal	Publication	http://purl.org/coar/resource_type/c_3e5a	A contribution to a journal denotes a work published in a journal. If applicable sub-terms sh	
20		Conference article		y	journal article	Publication	http://purl.org/coar/resource_type/c_6501	An article on a particular topic and published in a journal issue. (adapted from fabio)	
21	Contribution to Book/Report types								
22		Chapter (peer-reviewed)		y	book part	Publication	http://purl.org/coar/resource_type/c_3248	A defined chapter or section of a book, usually with a separate title or number.	
23		Chapter		y	book part	Publication	http://purl.org/coar/resource_type/c_3249	A defined chapter or section of a book, usually with a separate title or number.	
24		Entry for encyclopedia/		y	book part	Publication	http://purl.org/coar/resource_type/c_3250	A defined chapter or section of a book, usually with a separate title or number.	
25		Conference contribution		y	conference paper	Publication	http://purl.org/coar/resource_type/c_5794	A conference paper that is submitted to a conference and presented to the audience. The con	
26		Foreword/postscript		y	book part	Publication	http://purl.org/coar/resource_type/c_3252	A defined chapter or section of a book, usually with a separate title or number.	
27		Other chapter contribut		y	book part	Publication	http://purl.org/coar/resource_type/c_3253	A defined chapter or section of a book, usually with a separate title or number.	
28	Book/Report types								
29		Book		y	book	Publication	http://purl.org/coar/resource_type/c_2f33	A non-serial publication that is complete in one volume or a designated finite number of vol	
30		Anthology		y	book	Publication	http://purl.org/coar/resource_type/c_2f34	A non-serial publication that is complete in one volume or a designated finite number of vol	
31		Scholarly edition		y	book	Publication	http://purl.org/coar/resource_type/c_2f35	A non-serial publication that is complete in one volume or a designated finite number of vol	
32		Commissioned report		y	report	Publication	http://purl.org/coar/resource_type/c_93fc	A report is a separately published record of research findings, research still in progress, or of	
33		Other report		y	report	Publication	http://purl.org/coar/resource_type/c_93fc	A report is a separately published record of research findings, research still in progress, or of	

Publication/type mapping p2

Contribution to specialist publication types					
Article	y	journal article	Publication	http://purl.org/coar/resource_type/c_6501	An article on a particular topic and published in a journal issue. (adapted from fabio)
Featured article	y	journal article	Publication	http://purl.org/coar/resource_type/c_6501	An article on a particular topic and published in a journal issue. (adapted from fabio)
Book/Film/Article review	y	review	Publication	http://purl.org/coar/resource_type/c_efa0	A review of others' published work.
Editorial	y	editorial	Publication	http://purl.org/coar/resource_type/c_b239	A brief essay expressing the opinion or position of the chief editor(s) of a (academic) journal.
Letter	y	letter to editor	Publication	http://purl.org/coar/resource_type/c_2663	A letter to the editor is a letter sent to a periodical about issues of concern from its readers. (Adapted from fabio)
Special issue	y	periodical	Publication	http://purl.org/coar/resource_type/c_2664	A periodical is a publication issued on a regular and ongoing basis as a series of issues. (Adapted from fabio)
Non-textual form types					
Software	y	software	Product	http://purl.org/coar/resource_type/c_5ce6	A computer program in source code (text) or compiled form.
Data set/Database	y	dataset	Product	http://purl.org/coar/resource_type/c_ddb1	A collection of related facts and data encoded in a defined structure. (adapted from fabio; Datacite)
Digital or Visual Product	y	image	Product	http://purl.org/coar/resource_type/c_cf513	A visual representation other than text, including all types of moving image and still image.
Web publication/site	y	website	Product	http://purl.org/coar/resource_type/c_7ad9	A website, also written as web site or simply site, is a set of related web pages typically served by a web server.
Artefact	y	other	Product	http://purl.org/coar/resource_type/c_1843	A rest category which may cover text, interactive, sound, or image-based resources not explicitly covered by other categories.
Exhibition	y	other	Product	http://purl.org/coar/resource_type/c_1843	A rest category which may cover text, interactive, sound, or image-based resources not explicitly covered by other categories.
Performance	y	other	Product	http://purl.org/coar/resource_type/c_1843	A rest category which may cover text, interactive, sound, or image-based resources not explicitly covered by other categories.
Composition	y	musical composition	Product	http://purl.org/coar/resource_type/c_18cd	Musical composition can refer to an original piece of music, the structure of a musical piece, or a recorded static visual representation. This class of image includes diagrams, drawings, graphs, etc.
Design	y	Still image	Product	http://purl.org/coar/resource_type/c_ecc8	A recorded static visual representation. This class of image includes diagrams, drawings, graphs, etc.
Memorandum/exposition					
Memorandum	y	memorandum	Publication	http://purl.org/coar/resource_type/c_18wz	It is a note, document or other communication that helps the memory by recording events or information.
Question & Answer/hearing	y	other	Publication		
Contribution to memorandum/exposition					
Memorandum contribution	y	memorandum	Publication	http://purl.org/coar/resource_type/c_18wz	It is a note, document or other communication that helps the memory by recording events or information.
Question & Answer/hearing contribution	y	other	Publication		
Other contribution					
Other contribution	y	other	Publication	http://purl.org/coar/resource_type/c_1843	A rest category which may cover text, interactive, sound, or image-based resources not explicitly covered by other categories.
ALSO					
Dataset		dataset	Product	http://purl.org/coar/resource_type/c_ddb1	
Patent		patent	Patent	http://purl.org/coar/resource_type/c_15cd	

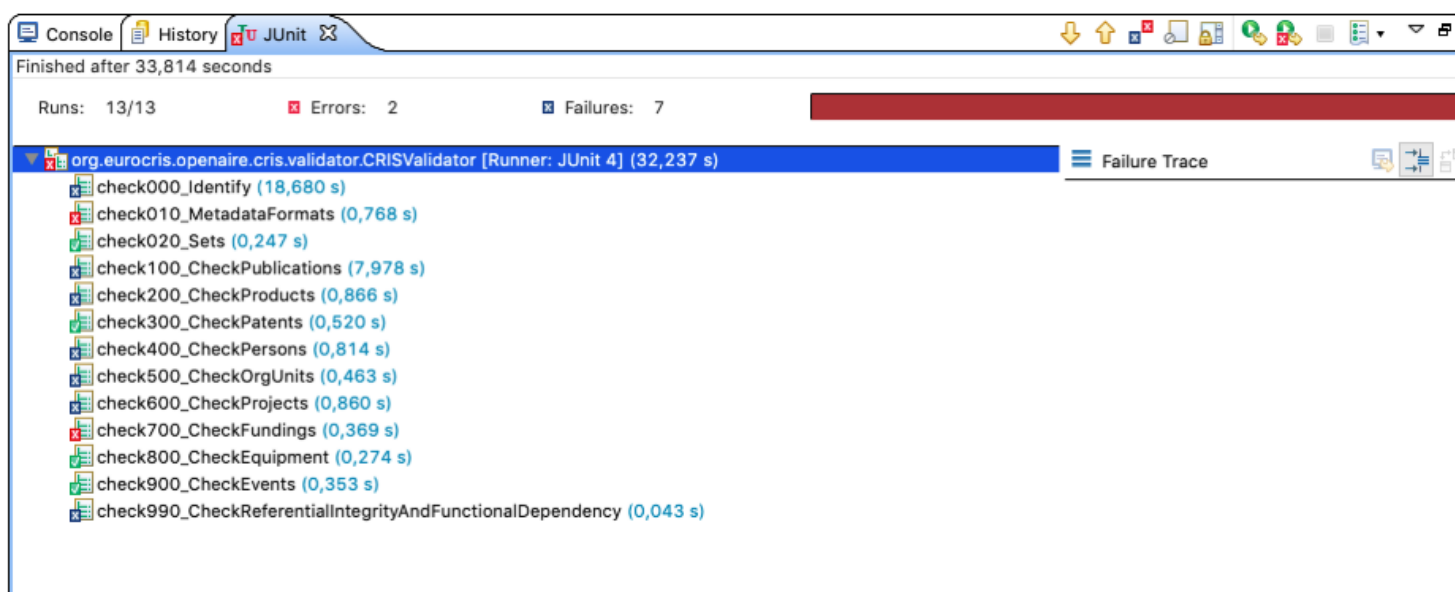
Implementation

- Internal Persons:
 - https://devel.atira.dk/eurocris/ws/oai?verb=ListRecords&metadataPrefix=oai_cerif_openaire&set=openaire_cris_persons
- Internal Organizations:
 - https://devel.atira.dk/eurocris/ws/oai?verb=ListRecords&metadataPrefix=oai_cerif_openaire&set=openaire_cris_orunits
- Research-Output:
 - https://devel.atira.dk/eurocris/ws/oai?verb=ListRecords&metadataPrefix=oai_cerif_openaire&set=openaire_cris_publications
- Projects:
 - https://devel.atira.dk/eurocris/ws/oai?verb=ListRecords&metadataPrefix=oai_cerif_openaire&set=openaire_cris_projects
- Events:
 - https://devel.atira.dk/eurocris/ws/oai?verb=ListRecords&metadataPrefix=oai_cerif_openaire&set=openaire_cris_events
- Equipment:
 - https://devel.atira.dk/eurocris/ws/oai?verb=ListRecords&metadataPrefix=oai_cerif_openaire&set=openaire_cris equipments

... and first validator results

Good to have the endpoint accessible.

When trying the prototype validator on that endpoint we see there are some issues:



In fact, you should be able to run the prototype validator (available from <https://github.com/jdvorak001/openaire-cris-validator>) internally. I'll be happy to provide clarifications of error messages where necessary.

Project timeline 2018/19

Feb	Pure dev version made available [5.11] with access to API, accompanying documentation and complete dataset for OpenAire mapping	Elsevier /Alex Kujath (AK)
Feb	Provide spreadsheet format of OpenAire CERIF-XML guidelines	euroCRIS/Jan Dvorak (JD)
Mar	Establish SLWG	euroCRIS/Anna Clements (AC)
Apr - Jun	Develop specification	SLWG
July	Implement specification	Elsevier
Aug/Sep	Feedback on implementation	SLWG
Oct/Dec	Test/Mature product	SLWG/Elsevier
Jan	Release 5.13.x	Elsevier

Implementation of the OpenAIRE CERIF-XML guidelines



How to make a CRIS OpenAIRE-compatible: the case of METIS: METIS2OpenAIRE case study

Ilaria Fava | 2018-08-22 | scholarly communication | Add Reply



Blogpost by Pablo de Castro, METIS2OpenAIRE project coordinator at euroCRIS

