

# Interoperability with Scholarly Infrastructure Services



DEC. 2019  
(REVISION)

Advancing Open Scholarship

## D7.1 – Interoperability with Scholarly Infrastructure Services

Version 1.1 – PUBLIC

This deliverable describes the OpenAIRE Interoperability with SHERPA/RoMEO and ScholeXplorer as Scholarly Infrastructure Services.



H2020-EINFRA-2017  
Grant Agreement 777541



# Document Description

## D7.1 – Interoperability with Scholarly Infrastructure Services

WP7 – Participatory Scholarly Communication	
WP participating organizations: CNR, ATHENA RC, UNIBI, UMINHO, EA, DARIAH ERIC, RBI	
Contractual Delivery Date: 10/2018	Actual Delivery Date: 01/2019
Nature: Report	Version: 1.1 refined
Public Deliverable	

### Preparation Slip

	Name	Organisation	Date
<b>From</b>	Andreas Czerniak	UNIBI	10/2018
	Jochen Schirrwagen	UNIBI	
	Alessia Bardi	CNR	
	Paolo Manghi	CNR	
<b>Edited by</b>	Andreas Czerniak	UNIBI	01/2019
<b>Reviewed by</b>	Lars Holm Nielsen	CERN	01/2019
	Alen Vodopivec	Rudjer Boskovic Institute	
	Azhar Hussain	Jisc	
<b>Approved by</b>	Paolo Manghi	CNR	01/2019
<b>For delivery</b>	Mike Chatzopoulos	UOA	01/2019

## Revision History

Issue	Item	Reason for Change	Author	Organization
V0.1	Draft version	Methodology and initial specifications; draft for review	Andreas Czerniak Jochen Schirrwagen	UNIBI
V1.0	Final Version	Reviewers Comments Implemented	Andreas Czerniak	UNIBI
V1.1	Refined Version	Revised based on the reviewers' comments.	Andreas Czerniak	UNIBI

## Table of Contents

1	INTERFACING WITH SCHOLARLY COMMUNICATION SERVICES	8
2	INTEROPERABILITY WITH SHERPA/ROMEO	9
2.1	Objective	9
2.2	Current status	9
2.3	Proposed Enhancements	11
	Example	13
	Scenario 1 - simple interlinkage	19
	Scenario 2 - simple interlink with information snippet	22
	Scenario 3 - advance interlink with scholarly information snippet	25
	Scenario 4 - connect to broker service	27
	Scenario 5 - improve advance scholarly interlinking	28
2.4	Integration in production	29
3	REPORT ON HOW TO INTEROPERATE WITH SCHOLARLY COMMUNICATION INFRASTRUCTURE	
-	DATA CITE	31
3.1	Current status	31
3.2	Proposed Enhancements	36
3.3	Integration in production	36
4	REFERENCES	37
5	FIGURES, TABLES, LISTINGS	38

---

# Disclaimer

---

This document contains description of the OpenAIRE-Advance project findings, work and products. Certain parts of it might be under partner Intellectual Property Right (IPR) rules so, prior to using its content please contact the consortium head for approval.

In case you believe that this document harms in any way IPR held by you as a person or as a representative of an entity, please do notify us immediately.

The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the OpenAIRE-Advance consortium and can in no way be taken to reflect the views of the European Union.

OpenAIRE-Advance is a project funded by the European Union (Grant Agreement No 777541).



## Acronyms

---

API	Application Programming Interface <sup>1</sup>
CC	Creative Commons <sup>2</sup>
DAI	Digital Author Identifier <sup>3</sup>
DOI	Digital Object Identifier <sup>4</sup>
EZB	Electronic Journals Library <sup>5</sup>
HTTPS	HyperText Transfer Protocol Secure
ISSN	International Standard Serial Number
JSON	JavaScript Object Notation
PID	Persistent Identifier
RDA	Research Data Alliance
REST	REpresentational State Transfer
SHERPA/RoMEO	Publisher copyright policies & self-archiving service from JISC
SHERPA/Juliet	database concerning funders' policies and more

---

<sup>1</sup> [https://en.wikipedia.org/wiki/Application\\_programming\\_interface](https://en.wikipedia.org/wiki/Application_programming_interface)

<sup>2</sup> <https://creativecommons.org/>

<sup>3</sup> [https://en.wikipedia.org/wiki/Digital\\_Author\\_Identifier](https://en.wikipedia.org/wiki/Digital_Author_Identifier)

<sup>4</sup> [https://en.wikipedia.org/wiki/Digital\\_object\\_identifier](https://en.wikipedia.org/wiki/Digital_object_identifier)

<sup>5</sup> <https://rzblx1.uni-regensburg.de/ezeit/>

## Publishable Summary

---

OpenAIRE and other scholarly communication infrastructure services mutually benefit from a fostered integration and utilization of their offered information. OpenAIRE collaborates with Jisc regarding SHERPA/RoMEO on the one hand and with DataCite regarding the ScholeExplorer on the other hand. These are two examples of efficient collaboration resulting in added value services and information enrichment.

This report discusses and describes several use cases and integration scenarios for both services. In the case of the SHERPA/RoMEO service, information on copyright policies and self-archiving options will inform authors of scientific publications in OpenAIRE and enrich bibliographic metadata together with relevant information, like funders archiving mandates and license information on national and consortial level. These measures are aimed to support Open Access to scientific publications.

The integration of the OpenAIRE ScholeExplorer service in DataCite ensures real-time updates of dataset to literature links, which is of benefit to third party services.



## 1 | INTERFACING WITH SCHOLARLY COMMUNICATION SERVICES

The purpose of the deliverable D7.1 is to present the scope and scenarios to enhance the interoperability of scholarly communication services with OpenAIRE. In particular we describe the interoperability of two selected scholarly infrastructure services, namely SHERPA/RoMEO with OpenAIRE and the OpenAIRE ScholeXplorer service with DataCite.

In the first part of the report we describe scenarios to integrate information of SHERPA/RoMEO in OpenAIRE. SHERPA is an award-winning service infrastructure in scholarly communication and is run by Jisc. The SHERPA infrastructure is operating several services that are of interest to be integrated into OpenAIRE.

First, we focus on SHERPA/RoMEO, which is a service for the publisher's copyright and archiving policies of journal articles and conference proceedings. The service aggregates and analyses publisher open access policies from around the world and provides summaries of self-archiving permissions and conditions of rights given to authors on a journal-by-journal basis.

SHERPA/JULIET is a service for research funders archiving mandates and guidelines.

In the second part of the report we focus on the integration of OpenAIRE ScholeXplorer<sup>6</sup> with DataCite metadata services.

---

<sup>6</sup> <https://dliservice.research-infrastructures.eu> , last accessed 2019-01-07

## 2 | INTEROPERABILITY WITH SHERPA/ROMEO

### 2.1 Objective

One of the key objectives of OpenAIRE is to support the implementation of open access mandates of research funding organizations as a measure to increase the share of open access publications in the total output of scientific publications. Authors of scientific publications are often unaware about the option of self-archiving their publications, which is further exacerbated by countless and inconsistent open access policies of publishers. As a consequence, authors often not use this opportunity and their publications stay behind paywalls.

SHERPA / RoMEO is a core service that collects and processes this information and is aimed to integrate with OpenAIRE. Information from this service may be used for the following purposes:

- to inform authors about self-archiving options,
- to facilitate the registration of hybrid journals in OpenAIRE and aggregate their metadata,
- to enrich the metadata of non-open access journal articles with SHERPA / RoMEO information in the OpenAIRE information graph.

### 2.2 Current status

“SHERPA/RoMEO<sup>7</sup>”, a service provided by Jisc, has been a preferred choice since many years. The RoMEO database contains information from 2566 publishers. 81% of them formally allow some kind of self-archiving<sup>8</sup> in form of a preprint or postprint, sometimes delayed by an embargo period. Actually, there are some developments that cover specific information about *publisher copyright policies & self-archiving information* but also policies resulting from alliance, national or consortium licenses which allow for secondary publication of scholarly works.

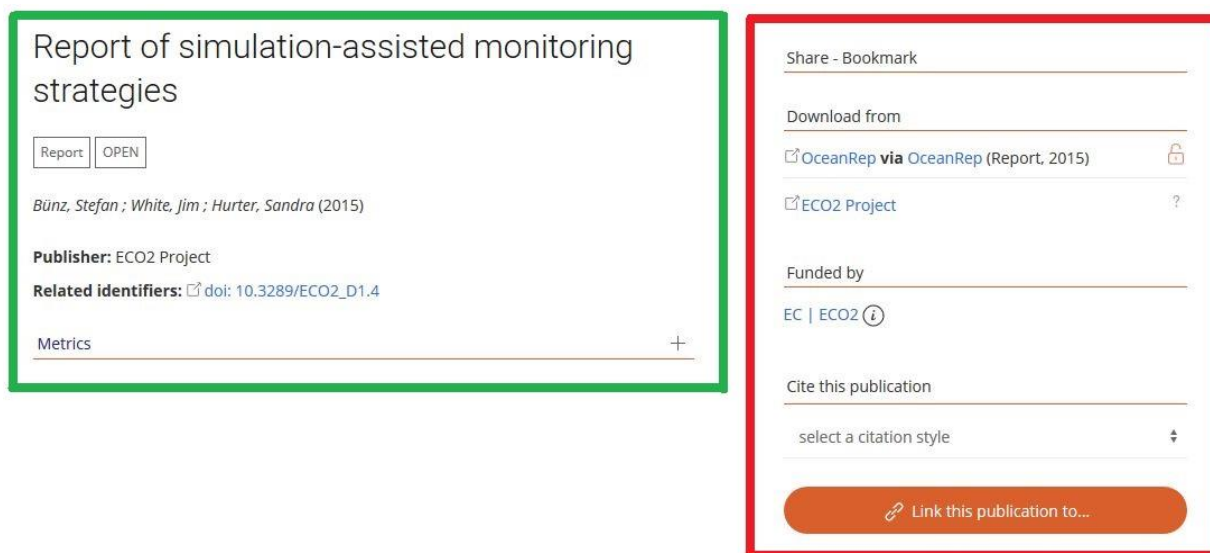
In the OpenAIRE discovery portal<sup>9</sup> an item (literature, dataset, software, ...) is displayed as depicted in Figure 1 and does not provide this information in the metadata.

---

<sup>7</sup> <http://www.sherpa.ac.uk/romeo> , last accessed 2018-10-12

<sup>8</sup> <http://www.sherpa.ac.uk/romeo/statistics.php> , last accessed 2019-01-04

<sup>9</sup> <https://explore.openaire.eu> , last accessed 2019-12-19



**FIGURE 1. ARTICLE METADATA AS SHOWN IN EXPLORE.OPENAIRE.EU**

On the left (green box), the metadata is presented by article name, resource type, access right, author(s), publisher, and related (persistent) identifier.

On the left side of Figure 1 (red box) options for sharing, download links, funding information, and citation style are presented. Provenance information about the origin of this record could be published here, too.

However, there is no information about the publishers' copyright policies and self-archiving options. For monitoring reasons, but also for the repository manager, it would be good to know under which conditions the article is published and which alternative options of self-archiving in Open Access exist, so that the author(s) can promote her/his work to a wider community.

The functional abstract diagram in Figure 2 describes this situation, according to [2].

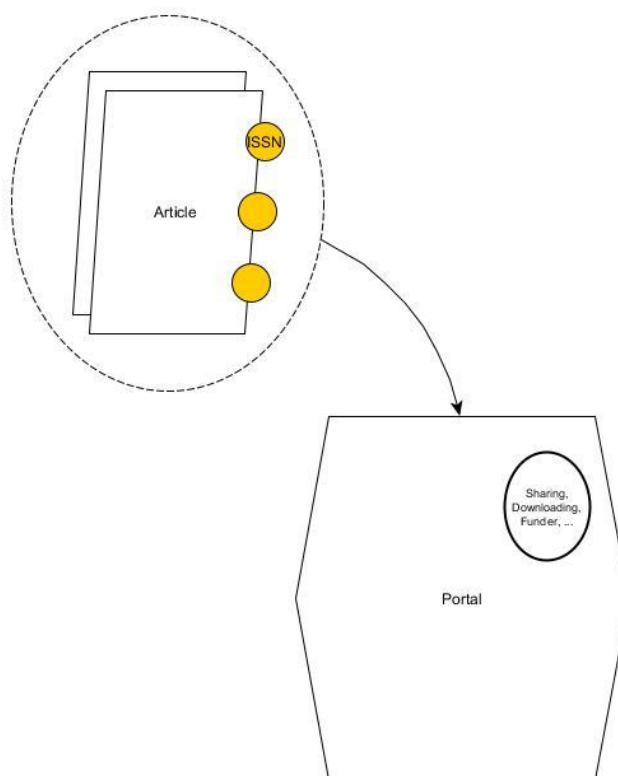


FIGURE 2. FUNCTIONAL ABSTRACT VIEW OF THE ARTICLE SHOWN AT EXPLORE.OPENAIRE.EU

## 2.3 Proposed Enhancements

In this section, we present the strategy to enrich bibliographic metadata on the portal (on article level) with the information about the publisher's copyright policies & self-archiving options. Additionally, we provide an outlook on the integration of further information systems (via web service interfaces) in order to enhance the information displayed on a detailed level.

In the following, we discuss and present some integration scenarios for its implementation.

Based on the “abstract view” as depicted in Figure 2, we develop new use cases to enhance the information on the article level.

The user interface of SHERPA/RoMEO can be accessed at <http://www.sherpa.ac.uk/romeo/search.php>. It returns the following information about the archiving policies for many journals:

RoMEO Colour	Archiving policy
<a href="#">Green</a>	Can archive pre-print <i>and</i> post-print or publisher's version/PDF
<a href="#">Blue</a>	Can archive post-print (ie final draft post-refereeing) or publisher's version/PDF
<a href="#">Yellow</a>	Can archive pre-print (ie pre-refereeing)
<a href="#">White</a>	Archiving not formally supported

FIGURE 3. TABLE OF ARCHIVING POLICY STATE (SCREENSHOT)

The complete list of “Definitions and Terms” are available at <http://www.sherpa.ac.uk/romeo/definitions.php?la=en&flDnum=|&mode=simple&version=-colours><sup>10</sup>. For this classification, we must know the ISSN attribute of an article.

---

<sup>10</sup> last access 2018-10-12

The information database can also be accessed via an Application Programming Interface (API). The documentation for this API can be found at <http://www.sherpa.ac.uk/romeo/apimanual.php?la=en&fIDnum=1>

The API is a RESTful interface; it is stateless and separates the concerns of client and server.

## Example

In this example, we explain how to retrieve information on self-archiving options of a journal in “SHERPA/RoMEO”.

In the next sections, we use the following ISSN “1444-1586” and the journal name “Geriatrics and Gerontology International” as an example.

Figure 4 presents information about this journal in SHERPA/RoMEO.



The screenshot displays the SHERPA/RoMEO website interface. At the top, the SHERPA/RoMEO logo is visible alongside the tagline "... opening access to research". Navigation links include Home, Search, Journals, Publishers, FAQ, Suggest, and About. The page title is "Search - Publisher copyright policies & self-archiving", and language options are provided in the top right corner.

The search results indicate that one journal was found for the query "geriatrics and gerontology international". The journal details are as follows:

- Journal:** [Geriatrics and Gerontology International](#) (ISSN: 1444-1586, ESSN: 1447-0594)
- RoMEO:** This is a RoMEO yellow journal
- Paid OA:** A paid open access option is available for this journal.
- Author's Pre-print:** ☒ author can archive pre-print (ie pre-refereeing)
- Author's Post-print:** ☒ subject to Restrictions below, author can archive post-print (ie final draft post-refereeing)
- Restrictions:**
  - 12 months embargo
- Publisher's Version/PDF:** ☒ author cannot archive publisher's version/PDF
- General Conditions:**
  - Some journals have separate policies, please check with each journal directly
  - On author's personal website, institutional repositories, arXiv, AgEcon, PhilPapers, PubMed Central, RePEc or Social Science Research Network
  - Author's pre-print may not be updated with Publisher's Version/PDF
  - Author's pre-print must acknowledge acceptance for publication
  - Non-Commercial
  - Publisher's version/PDF cannot be used
  - Publisher source must be acknowledged with citation
  - Must link to publisher version with set statement (see policy)
  - If OnlineOpen is available, BBSRC, EPSRC, MRC, NERC and STFC authors, may self-archive after 12 months
- Mandated OA:** (Awaiting information)
- Paid Open Access:** [OnlineOpen](#)
- Notes:**
  - Publisher last contacted on 07/08/2014
- Copyright:** [Self-archiving - Authors Compliance Tool - Funder Policies](#)
- Updated:** 24-Nov-2017 - [Suggest an update for this record](#)
- Link to this page:** <http://sherpa.ac.uk/romeo/issn/1444-1586/>
- Published by:** [Wiley: 12 months](#) [Commercial Publisher] - [Yellow Policies in RoMEO](#)
- For:** [Japan Geriatrics Society](#) [Copyright Holder]
- Guidance:** Please see the list of [Publisher Categories in RoMEO](#) for guidance on interpreting the priority of multiple publishers.

A disclaimer at the bottom states: "These summaries are for the journal's default policies, and changes or exceptions can often be negotiated by authors. All information is correct to the best of our knowledge but should not be relied upon for legal advice."

FIGURE 4. INFORMATION ABOUT A JOURNAL IN SHERPA/RoMEO

By the help of the ISSN, we can look up the same information also via the API. The request for our example is:

<http://www.sherpa.ac.uk/romeo/api29.php?issn=1444-1586>

This request can also be transformed into a REST URL:

<http://www.sherpa.ac.uk/romeo/api29/issn/1444-1586/>

The complete response of the query above from SHERPA/RoMEO is shown in Listing 1:

```
<!DOCTYPE romeoapi SYSTEM "http://www.sherpa.ac.uk/romeo/romeoapi293.dtd">
<romeoapi version="2.9.9">
  <header>
    <parameters>
      <parameter source="specified">
        <parametername>issn</parametername>
        <parametervalue>1444-1586</parametervalue>
      </parameter>
      <parameter source="default">
        <parametername>showfunder</parametername>
        <parametervalue>none</parametervalue>
      </parameter>
      <parameter source="default">
        <parametername>fIDnum</parametername>
        <parametervalue>|</parametervalue>
      </parameter>
      <parameter source="default">
        <parametername>la</parametername>
        <parametervalue>en</parametervalue>
      </parameter>
    </parameters>
    <numhits>1</numhits>
    <apicontrol>journal</apicontrol>
    <outcome>singleJournal</outcome>
    <message/>
    <licence>SHERPA/RoMEO data is available for re-use under a Creative
    Commons Attribution-Non-Commercial-Share Alike 2.5 licence. For more details,
    please see the 'conditions for re-use' at the &lt;licenceurl /&gt;; and linked-
    to from the SHERPA/RoMEO home page.</licence>
    <licenceurl>http://www.sherpa.ac.uk/romeoreuse.html</licenceurl>
    <disclaimer>All SHERPA/RoMEO information is correct to the best of our
    knowledge but should not be relied upon for legal advice. SHERPA cannot be held
    responsible for the re-use of RoMEO data, or for alternative interpretations
    which are derived from this information.</disclaimer>
    <timestamp>13-Dec-2018:12:31:09</timestamp>
  </header>
  <journals>
    <journal>
      <jtitle>Geriatrics and Gerontology International</jtitle>
      <issn>1444-1586</issn>
      <zetocpub>Wiley: 12 months</zetocpub>
      <romeopub>Wiley: 12 months</romeopub>
    </journal>
  </journals>
```

```

<publishers>
  <publisher id="2050" parentid="580">
    <name>Wiley</name>
    <alias>12 months</alias>
    <homeurl>http://eu.wiley.com/WileyCDA/</homeurl>
    <preprints>
      <prearchiving>can</prearchiving>
      <prerestrictions/>
    </preprints>
    <postprints>
      <postarchiving>restricted</postarchiving>
      <postrestrictions>
        <postrestriction>&lt;num&gt;12&lt;/num&gt; &lt;period
units="month"&gt;months&lt;/period&gt; embargo</postrestriction>
      </postrestrictions>
    </postprints>
    <pdfversion>
      <pdfarchiving>cannot</pdfarchiving>
      <pdfrestrictions/>
    </pdfversion>
    <conditions>
      <condition>Some journals have separate policies, please check with each
journal directly</condition>
      <condition>On author's personal website, institutional repositories,
arXiv, AgEcon, PhilPapers, PubMed Central, RePEc or Social Science Research
Network</condition>
      <condition>Author's pre-print may not be updated with Publisher's
Version/PDF</condition>
      <condition>Author's pre-print must acknowledge acceptance for
publication</condition>
      <condition>Non-Commercial</condition>
      <condition>Publisher's version/PDF cannot be used</condition>
      <condition>Publisher source must be acknowledged with
citation</condition>
      <condition>Must link to publisher version with set statement (see
policy)</condition>
      <condition>If OnlineOpen is available, BBSRC, EPSRC, MRC, NERC and STFC
authors, may self-archive after 12 months</condition>
      <condition>Publisher last contacted on 07/08/2014</condition>
    </conditions>
    <mandates/>
    <paidaccess>
      <paidaccessurl>https://authorservices.wiley.com/author-
resources/Journal-Authors/licensing-open-access/open-
access/onlineopen.html</paidaccessurl>
      <paidaccessname>OnlineOpen</paidaccessname>
      <paidaccessnotes>A paid open access option is available for this
journal.</paidaccessnotes>
    </paidaccess>
    <copyrightlinks>
      <copyrightlink>
        <copyrightlinktext>Self-archiving</copyrightlinktext>
        <copyrightlinkurl>http://olabout.wiley.com/WileyCDA/Section/id-
828039.html</copyrightlinkurl>
      </copyrightlink>
      <copyrightlink>
        <copyrightlinktext>Authors Compliance Tool</copyrightlinktext>
        <copyrightlinkurl>http://olabout.wiley.com/WileyCDA/Section/id-
828038.html</copyrightlinkurl>
      </copyrightlink>
    </copyrightlinks>
  </publisher>
</publishers>

```



```

    <copyrightlink>
    <copyrightlinktext>Funder Policies</copyrightlinktext>
    <copyrightlinkurl>https://authorservices.wiley.com/author-
resources/Journal-Authors/licensing-open-access/open-access/funder-
agreements.html</copyrightlinkurl>
    </copyrightlink>
    </copyrightlinks>
    <romeocolour>yellow</romeocolour>
    <dateadded>2014-08-08 15:08:35</dateadded>
    <dateupdated>2017-11-24 10:22:27</dateupdated>
    </publisher>
  </publishers>
</romeoapi>

```

**LISTING 1. RESULT OF THE API QUERY CALL.**

Figure 5 shows the corresponding section about the restrictions to make the article open access as they are displayed in the web browser:

<b>Author's Pre-print:</b>	✓ author can archive pre-print (ie pre-refereeing)
<b>Author's Post-print:</b>	⚠ subject to Restrictions below, author can archive post-print (ie final draft post-refereeing)
<b>Restrictions:</b>	<ul style="list-style-type: none"> <li>• 12 months embargo</li> </ul>
<b>Publisher's Version/PDF:</b>	✗ author cannot archive publisher's version/PDF

**FIGURE 5. RESULTS OF A RESTRICTION OF THE SEARCH IN A BROWSER (PARTIAL).**

Listing 2 shows the same response snippet from the RESTful interface.

```

<preprints>
  <prearchiving>can</prearchiving>
  <prerestrictions/>
</preprints>
<postprints>
  <postarchiving>restricted</postarchiving>
  <postrestrictions>
    <postrestriction>&lt;num>12</num> &lt;period
units="month">months</period> embargo</postrestriction>
  </postrestrictions>
</postprints>
<pdfversion>
  <pdfarchiving>cannot</pdfarchiving>
  <pdfrestrictions/>
</pdfversion>

```

**LISTING 2. RESTRICTION RESULT SET IN THE API QUERY CALL (PARTIAL).**

The next interesting section is the “general conditions” from Figure 4. A closer look is presented in Figure 6.

General Conditions:	<ul style="list-style-type: none"> <li>• Some journals have separate policies, please check with each journal directly</li> <li>• On author's personal website, institutional repositories, arXiv, AgEcon, PhilPapers, PubMed Central, RePEc or Social Science Research Network</li> <li>• Author's pre-print may not be updated with Publisher's Version/PDF</li> <li>• Author's pre-print must acknowledge acceptance for publication</li> <li>• Non-Commercial</li> <li>• Publisher's version/PDF cannot be used</li> <li>• Publisher source must be acknowledged with citation</li> <li>• Must link to publisher version with set statement (see policy)</li> <li>• If OnlineOpen is available, BBSRC, EPSRC, MRC, NERC and STFC authors, may self-archive after 12 months</li> </ul>
---------------------	---

FIGURE 6. GENERAL CONDITIONS OF RESTRICTIONS AS RESULTS OF THE SEARCH IN A BROWSER.

The information provided is an overview of the *use and re-use conditions* of the published article in text form. The same information snippet retrieved via the API layer is shown in listing 3. The *<condition>* elements are free-text fields with some semantic notions.

```
<conditions>
  <condition>Some journals have separate policies, please check with
each journal directly</condition>
  <condition>On author's personal website, institutional
repositories, arXiv, AgEcon, PhilPapers, PubMed Central, RePEc or
Social Science Research Network</condition>
  <condition>Author's pre-print may not be updated with Publisher's
Version/PDF</condition>
  <condition>Author's pre-print must acknowledge acceptance for
publication</condition>
  <condition>Non-Commercial</condition>
  <condition>Publisher's version/PDF cannot be used</condition>
  <condition>Publisher source must be acknowledged with
citation</condition>
  <condition>Must link to publisher version with set statement (see
policy)</condition>
  <condition>If OnlineOpen is available, BBSRC, EPSRC, MRC, NERC and
STFC authors, may self-archive after 12 months</condition>
  <condition>Publisher last contacted on 07/08/2014</condition>
</conditions>
```

LISTING 3. CONDITION RESULT SET IN THE API QUERY CALL (PARTIAL).

As a result, the service may enrich the bibliographic metadata with specific information about the publisher’s conditions. It informs the author(s) and repository managers so that they can decide to deposit an open access version of an article.

The information in the SHERPA/RoMEO database itself is subject to the CC-BY-NC-ND<sup>11</sup> license. It is thus free to share - copy and redistribute the material in any medium or format, under citation, and for non-commercial purposes.

A complementary service to “SHERPA/RoMEO” is “SHERPA/Juliet”.

“SHERPA/Juliet” enables researchers and librarians to see funders' conditions to make research output open access. The service is accessible via a web browser or from a machine-readable API.

The URL of the service is

<http://v2.sherpa.ac.uk/juliet/>

The RESTful API service is

<https://v2.sherpa.ac.uk/cgi/retrieve>

### Examples

- `/cgi/retrieve?item-type=funder&api-key=YOUR-API-KEY&format=Json&limit=5&offset=10`
- `/cgi/retrieve?item-type=funder&api-key=YOUR-API-KEY&format=Json&filter=[["country","equals","ca"]]`

In the following, we describe use case scenarios/models[1] to use the previously presented interface with its information in order to:

- inform authors about self-archiving options including publisher embargo periods in the OpenAIRE portal,
- collect information data on hybrid journals and to support their registration in OpenAIRE for metadata harvesting, and to
- enrich metadata of (non-)open-access journal contributions in the OpenAIRE information space.

By integration of SHERPA/RoMEO and SHERPA/Juliet information in OpenAIRE, authors get informed about their compliance with funder's Open Access policies and mandates. Furthermore, the Open Access status of a journal registered in OpenAIRE can be determined and finally, an open access version of a paywalled article can potentially be made available.

---

<sup>11</sup> <https://creativecommons.org/licenses/by-nc-nd/4.0/>

## Scenario 1 - simple interlinkage

In this first scenario, the article metadata is enriched with an information link to the SHERPA/RoMEO database and its interface.

To establish our enhancement, the following steps are required:

- aggregating/collecting workflow, which collected information from SHERPA/RoMEO as link and
- serialize the link for the specific ISSN in an information area to
- join/inject in a component both, the newly collected information from SHERPA/RoMEO is a link to the specific information with the article metadata.

Afterward, the information context of a certain article and its representation in [explore.openaire.eu](http://explore.openaire.eu) makes it possible to provide the link to the self-archiving and embargo information under the text label of SHERPA/RoMEO. These are conceptually shown in Figure 8.

The aggregation in this scenario covers the availability of the specific ISSN of a journal in the SHERPA/RoMEO database and serializes the connection to the information space.

If an article is published with an ISSN, and this ISSN is found in the information space, then the link text to SHERPA/RoMEO is shown as e.g.

Publisher copyright policies & self-archiving

While clicking on the link of the anchor text would open a separate window:

<http://www.sherpa.ac.uk/romeo/search.php?issn=1364-453X&la=en>

In the portal, this could be integrated on the right side (red box of figure 1). Figure 7 is shown a mockup on the article level.

## Report of simulation-assisted monitoring strategies

Report OPEN

Bünz, Stefan ; White, Jim ; Hurter, Sandra (2015)

**Publisher:** ECO2 Project

**Related identifiers:**  doi: 10.3289/ECO2\_D1.4

Metrics

Share - Bookmark

Download from

 OceanRep via OceanRep (Report, 2015) 

 ECO2 Project ?

Funded by

EC | ECO2 

Cite this publication

select a citation style

Selfarchiving and embargo information  
[link to Sherpa/RoMEO](#)


 Link this publication to...

FIGURE 7. MOCKUP OF SCENARIO 1

The code snippet in HTML looks like:

```
<a href="http://www.sherpa.ac.uk/romeo/search.php?issn=1364-453X&la=en"
target="_blank">link to SHERPA/RoMEO</a>
```

A special feature could be to open the link in the user-specific language if it is known.

The result of this scenario is an automatically generated information snippet that integrates the reference to SHERPA/RoMEO. If the information snippet contains information on "embargo periods", this information could also be introduced in the additional field on the OpenAIRE Explore portal.

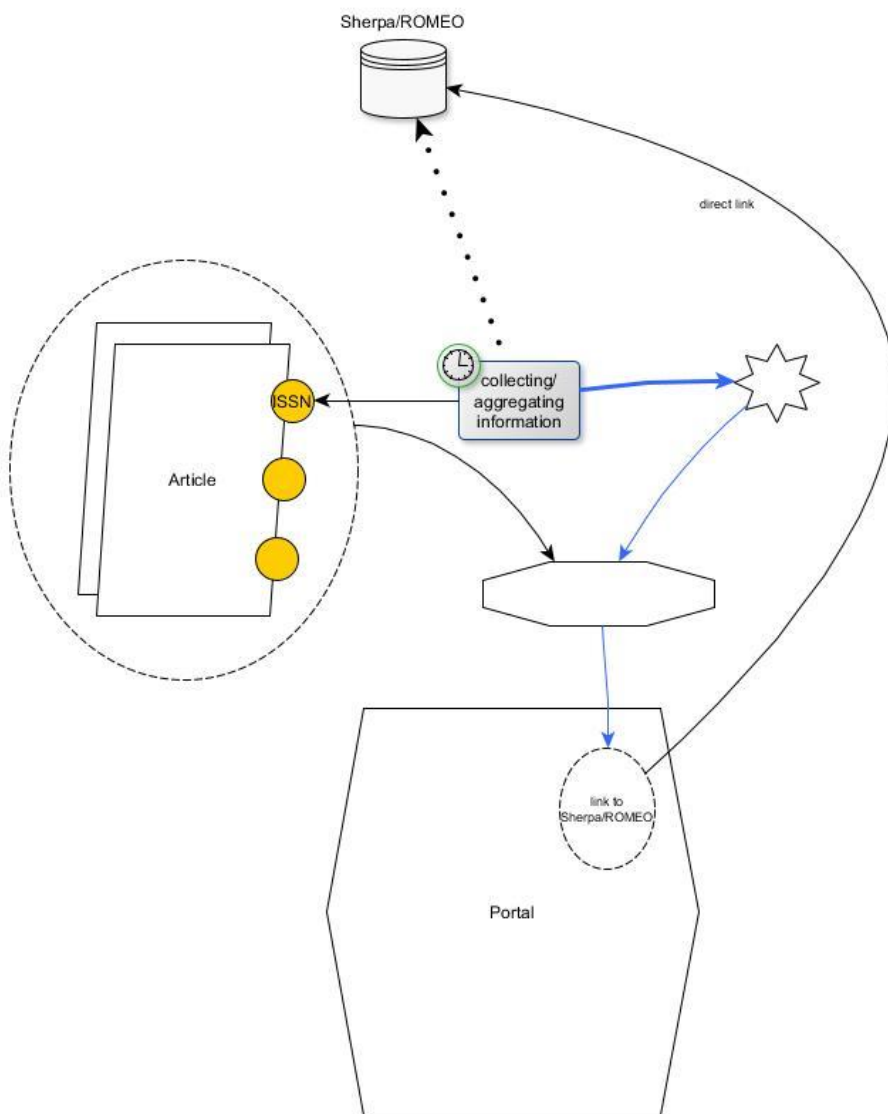


FIGURE 8. FUNCTIONAL ABSTRACT VIEW OF SCENARIO 1: SIMPLE INTERLINKAGE

## Scenario 2 - simple interlink with information snippet

This scenario is an enhancement of the “scenario 1 interlinkage”:

- aggregate/collect the specific information from “SHERPA/RoMEO” using the ISSN parameter,
- store the collected information in a separate information space displayed as a star in figure 9,
- join/inject this collected information with the identification of the article, e.g. DOI, combined with the ISSN and the specific information from SHERPA, depends on the time range in how this article is published. This is due to the fact that the policies may change over time.

The result of this newly collected information is presented on the [explore.openaire.eu](https://explore.openaire.eu) website in the same way as the first information extension is shown in Figure 8. The difference is that the time component is now included and the correct information of publishing date and information from SHERPA can be displayed. Figure 9 illustrates this concept.

This extension is not visible to the scientist, but is significant for the trustworthiness of the OpenAIRE Portal:

Publisher copyright policies & self-archiving  
*with compact description of figure 4/table 2, e.g.*

- author can archive pre-prints
- author can archive post-prints **after 12 month of embargo**
- author can't archive the publisher version

This text block can be enriched with the link to the original source. It can also contain more general conditions like the following:

- Author's pre-print on pre-print servers such as arXiv.org
- Author's post-print on author's personal website immediately
- Author's post-print on any open access repository after 12 months after publication
- Publisher's version/PDF cannot be used
- Published source must be acknowledged
- Must link to publisher version
- Set phrase to accompany link to published version (see policy)
- Articles in some journals can be made Open Access on payment of additional charge



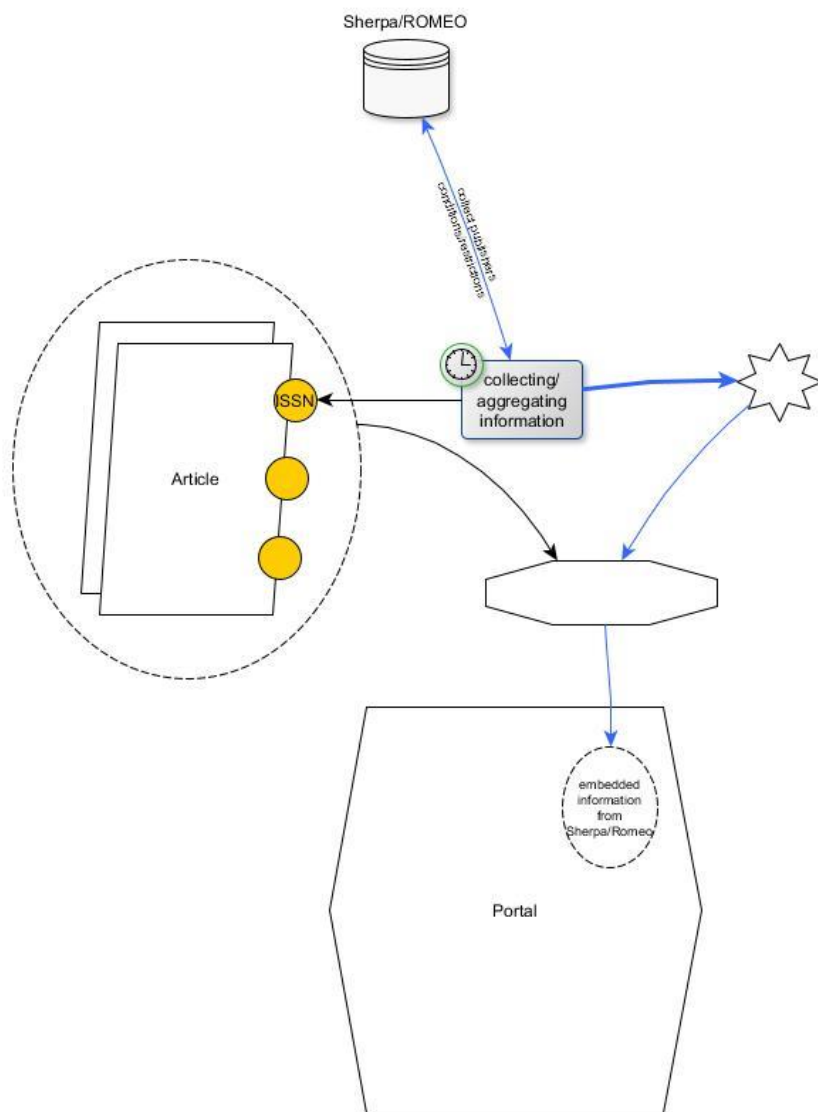


FIGURE 9. FUNCTIONAL ABSTRACT VIEW OF SCENARIO 2: SIMPLE INTERLINK WITH INFORMATION SNIPPET

## Scenario 3 - advance interlink with scholarly information snippet

This scenario enhances scenario 2. The previous scenario collects information about the publisher's copyright and archiving policies. In this case, we extend the information from SHERPA/Juliet with funder archiving mandates and guidelines. This scenario is one of the steps to test the objectives of information from the archiving mandates and guidelines of the national funders.

The following objectives are shown in figure 9:

- aggregate/collect the information from SHERPA/RoMEO on the ISSN identification for publisher copyright & archiving policy,
- aggregate/collect the information piece from SHERPA/Juliet on the ISSN identification for funders archiving mandates and guidelines,
- store these information sets in a separate storage area, displayed as a star,
- join/inject this collected information with the identification of the article, e.g. DOI, combined with the ISSN and the specific information from SHERPA and the ISSN, depends on the time range in how this article is published. This is due to the fact that the policies may change over time,
- in particular, the evaluation of the possibility of integrating "national funding agencies" with their restrictions on the publication of articles. This can be based on the DOI or a combination of other PIDs,
- identify hybrid journals to support their registration in OpenAIRE.

The goal of this scenario is to combine different sources of information that are useful for authors and present these in an extended field, as shown in Figure 7.

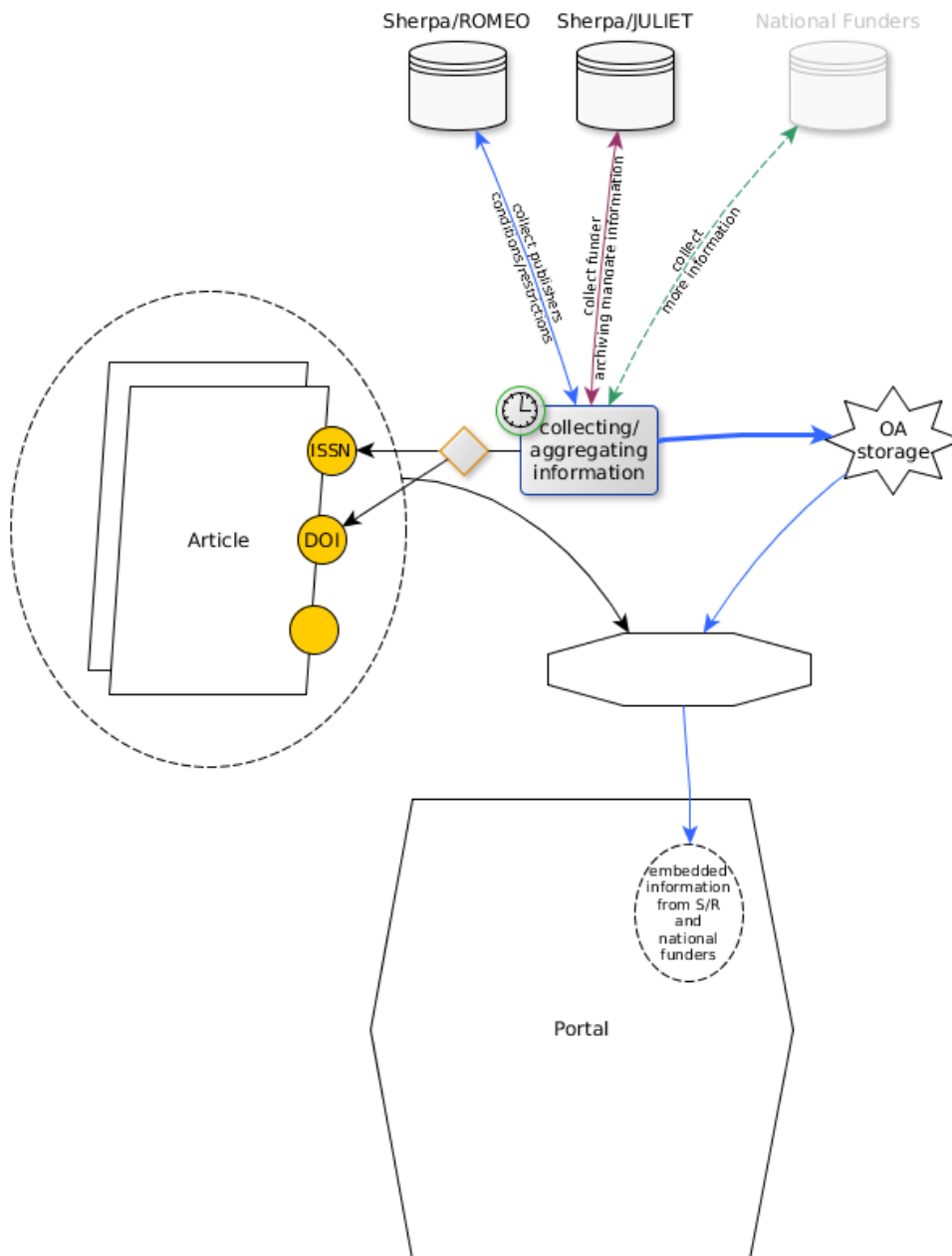


FIGURE 10. FUNCTIONAL ABSTRACT VIEW OF SCENARIO 3: ADVANCE INTERLINK WITH SCHOLARLY INFORMATION SNIPPET

## Scenario 4 - connect to broker service

This scenario will provide the information snippets from scenario 3 as a piece of topical information to the *Information Broker Service* of OpenAIRE.

The OpenAIRE *Information Broker Service* notifies interested (subscribed) repositories and CRIS systems about publication metadata information available in OpenAIRE.

It will identify records of publications that are (potentially) of interest to subscribed repositories and CRIS systems and will send the information about these records to the repository and CRIS managers. As examples: the publisher article is closed access, but the pre-/postprint can be open access, or the embargo date is already expired.

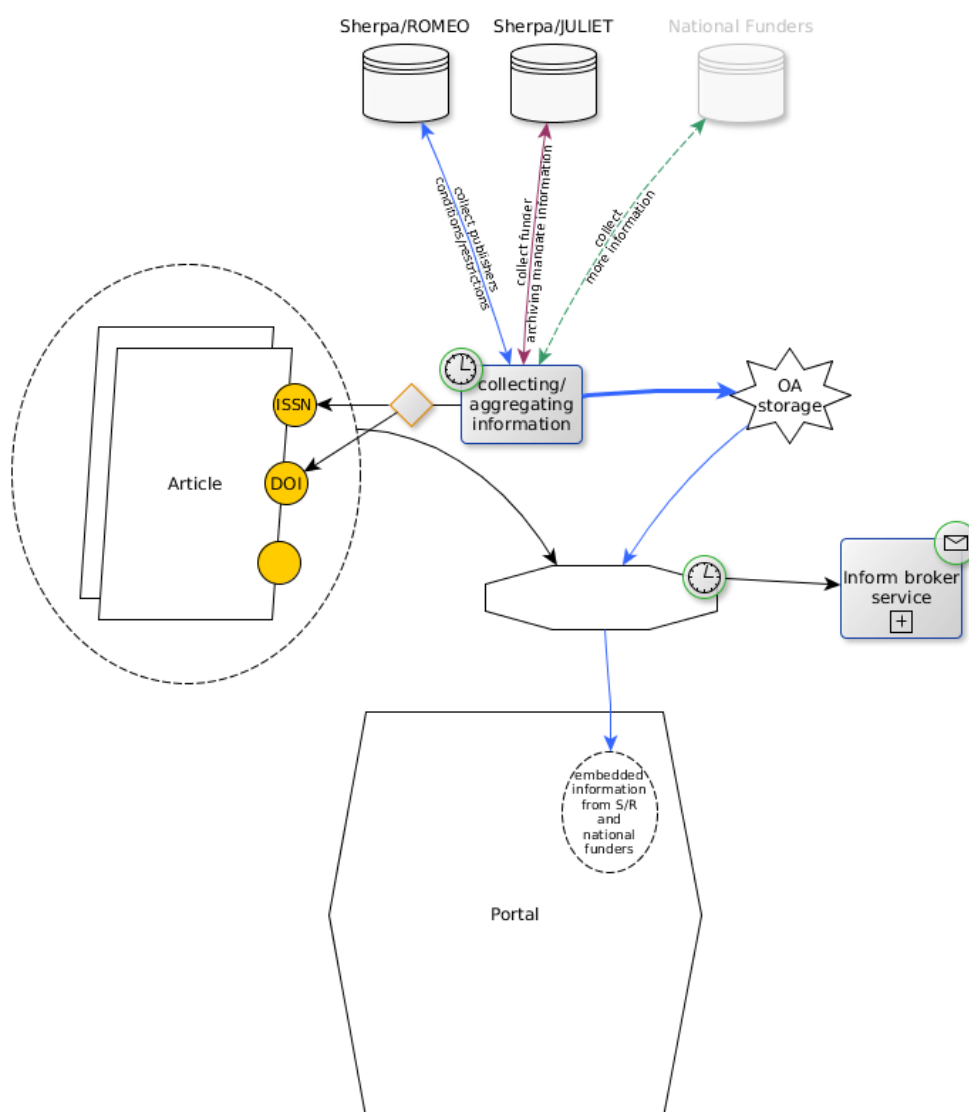


FIGURE 11. FUNCTIONAL ABSTRACT VIEW OF SCENARIO 4: CONNECT TO BROKER SERVICE

## Scenario 5 - improve advance scholarly interlinking

This last scenario enhances the view and integrates all cases with more than one national resource for archiving mandates, guidelines and licenses (national or consortial regulations).

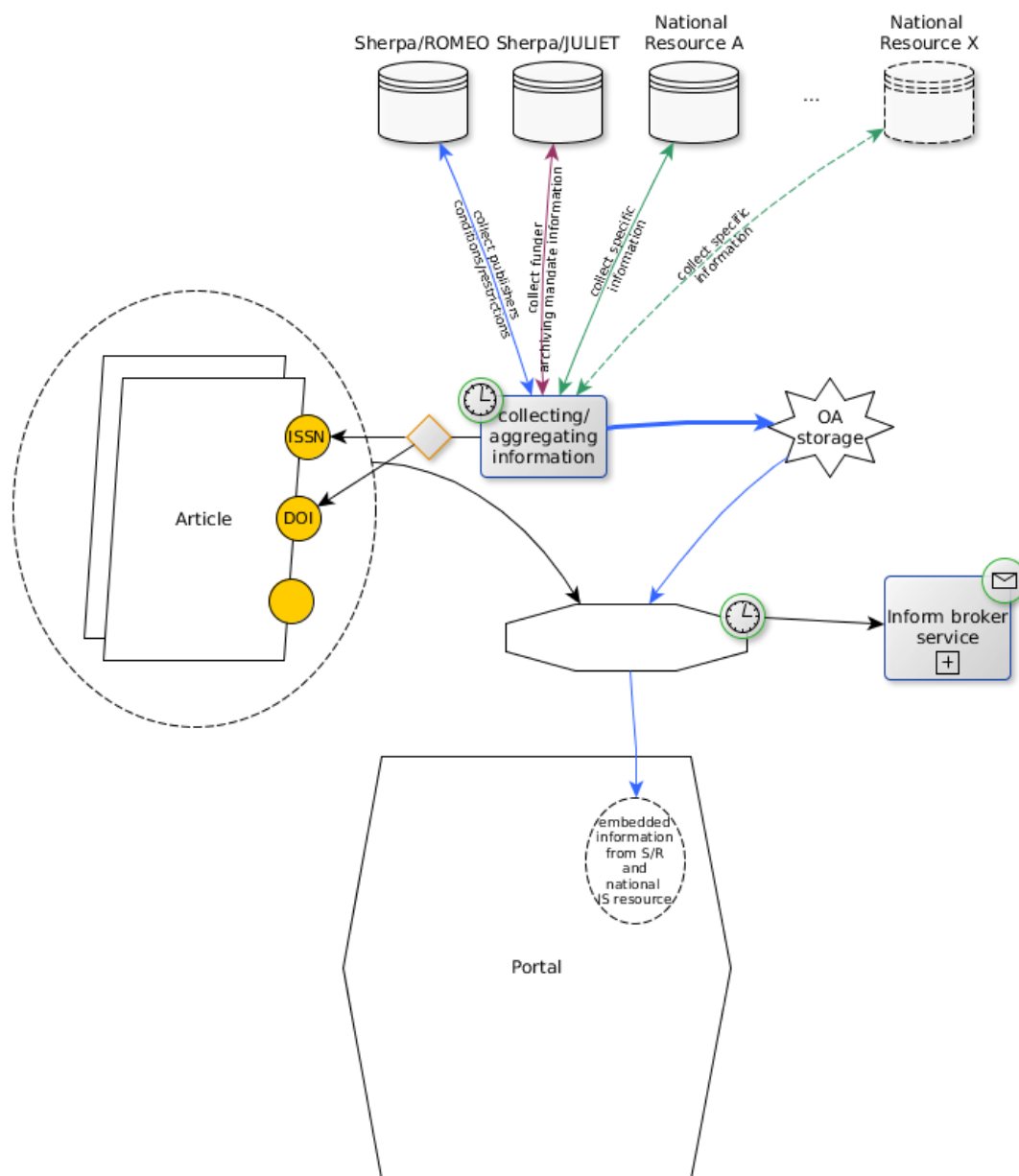


FIGURE 12. FUNCTIONAL ABSTRACT VIEW OF SCENARIO 5: IMPROVE ADVANCE SCHOLARLY INTERLINKING

With this last scenario, all required components to collect information are available and can be made available in the OpenAIRE explore portal and information graph.

Every single article is now enriched with information on self-archiving options and thus provides added value for the author, as well as for the repository manager.

## 2.4 Integration in production

We suggest the development of a prototype application to exploit the API information to be used in OpenAIRE.

The prototype application for the first three scenarios can be done by a single application for collecting the information from “SHERPA/RoMEO” and “SHERPA/Juliet”. Based on this, the modules of the single application can be integrated. To this aim, the application should be part of the workflow orchestration of OpenAIRE services which requires further coordination with teams in other work packages.

Furthermore, the representation of the publisher's copyright information and self-archiving policies in the OpenAIRE portal needs to be discussed.

The integration of the previously listed data sources in OpenAIRE will require the generation of a consolidated database about journal titles and their ISSN attributes. Quality curation mechanisms and continuous monitoring functions of the aggregation subsystem shall be used to create, update or discontinue items, including their time periods, during which the information is valid. Here, the time domain refers to the area in which the information is valid, since (national) policies may change over time.

One of the essential issues for a quality control mechanism is to know which records are changed, which records are added.

For the SHERPA/RoMEO service, the service offers a field with date and time on which the publisher's record was last edited/updated in the SHERPA/RoMEO database. It can be used within the quality control mechanism.

The field (XML) `<dateupdated>` hold these information

```
<dateupdated>2017-11-24 10:22:27</dateupdated>
```

(example from Listing 1, page 15).

In addition, the SHERPA/RoMEO API can be used for harvesting in incremental mode, e.g. to query for new or modified entries using a specific date.

The call for the API is:

<http://www.sherpa.ac.uk/romeo/api29.php?pdate=<specified date>>

However, there is a limitation. The query doesn't indicate which field is changed or which record was added. A separate procedure is required for curation in case no information on the removal, deletion, or discontinuation of records is found.

## 3 | REPORT ON HOW TO INTEROPERATE WITH SCHOLARLY COMMUNICATION INFRASTRUCTURE - DATACITE

### 3.1 Current status

The ScholeXplorer service, operated by OpenAIRE, has now reached almost a year in production and collects scholarly links article-dataset and dataset-dataset from:

- CrossRef's EventData Service: exposes article-dataset links made available by publishers as Scholix records;
- DataCite: exposes article-dataset links and dataset-dataset links made available by data centers and repositories members of DataCite, via DataCite records;
- OpenAIRE graph: exposes article-dataset links inferred by mining article PDFs in search of DOIs of datasets and accession numbers, as OpenAIRE compliant records

OpenAIRE ScholeXplorer has integrated with DataCite metadata services in order to keep its index up-to-date every hour. The frequency was set in agreement with the consumers of ScholeXplorer and with DataCite, in order to ensure the availability of links from datasets to articles as soon as possible after they are published by the researchers at the data center sites. Thanks to the almost real-time integration, third-party services like Scopus, Zenodo, and other repositories are today relying on ScholeXplorer to identify links to literature on demand. The ScholeXplorer API are described at <http://scholexplorer.openaire.eu/#/api>, while a technical documentation on the API is available at <http://api.scholexplorer.openaire.eu/v2/ui/>. The available API methods are listed in Figure 1, which shows a screenshot of the API technical documentation.

#### Scholexplorer API 2.0 2.0.0 OAS3

[/v2/openapi.json](#)

scholexplorer API version 2.0

##### LinkProvider : Operation related to the Link Provider

**GET** [/v2/LinkProvider](#) Get All Link Providers

##### LinkPublisher : Operation related to the Link Publisher

**GET** [/v2/LinkPublisher/inSource](#) Get All Publishers that provide source object

**GET** [/v2/LinkPublisher/inTarget](#) Get All Publishers that provide target object

##### Links : Operation related to the Scholix Links

**GET** [/v2/Links](#) Get Scholix Links

Figure 1 ScholeXplorer API documentation



- `/v2/LinkProvider` returns the names of providers of links and, for each of them, the number of collected links. The method takes an optional argument in input that allows to filter the providers by name. The schema of the response is defined in Figure 2

```

Response_Linkprovider ▾ {
  title: Response_Linkprovider
  LinkProvider ▾ {
    name                string
                        title: The name of the Provider that provides the links
                        maxLength: 300
    totalRelationships integer
                        title: The number of links that It provides
  }
}

```

*Figure 2 Response schema for the API method `/v2/LinkProvider`*

- `/v2/LinkPublisher/inSource` returns the names of providers of objects that have an outgoing relationship. For each provider in the response, the number of links outgoing from its object is also returned. The method takes an optional argument in input that allows filtering the providers by name. The format of the response is the same as the one for `/v2/LinkProvider` shown in Figure 2.
- `/v2/LinkPublisher/inTarget` returns the names of providers of objects that have an incoming relationship. For each provider in the response, the number of incoming links to its object is also returned. The method takes an optional argument in input that allows filtering the providers by name. The format of the response is the same as the one for `/v2/LinkProvider` shown in Figure 2.
- `/v2/Links` returns a list of Scholix links. The method allows to specify a set of parameters, as described in Figure 3. The response (Figure 4) includes information about the number of links matching the request and a list of links in Scholix format, compliant to the schema

available at <https://github.com/scholix/schema/tree/master/json/v3/schema.json>.

GET /v2/Links Get Scholix Links	
Return a List of scholix Links, this method <b>require one of the following parameters</b> (sourcePid, targetPid, sourcePublisher, \targetPublisher, linkProvider) all parameters can be combined	
<b>Parameters</b>	
Name	Description
linkProvider string (query)	Filter Scholix relationships collected from a LinkProvider  linkProvider - Filter Scholix relationships colle
targetPid string (query)	Filter Scholix relationships having a target pid  targetPid - Filter Scholix relationships having
targetPublisher string (query)	Filter Scholix relationships having a target published in a Publisher named targetPublisher  targetPublisher - Filter Scholix relationships h
targetPidType string (query)	Filter Scholix relationships having a target pid type  targetPidType - Filter Scholix relationships he
sourcePid string (query)	Filter Scholix relationships having a source pid  sourcePid - Filter Scholix relationships havin
sourcePublisher string (query)	Filter Scholix relationships having a source published in a Publisher named sourcePublisher  sourcePublisher - Filter Scholix relationships
sourcePidType string (query)	Filter Scholix relationships having a source pid type  sourcePidType - Filter Scholix relationships h
harvestedAfter string (query)	Filter scholix Links having collected after this date  harvestedAfter - Filter scholix Links having cc
page integer (query)	select page of result  page - select page of result

Figure 3 Parameters for the PAI method /v2/Links

```

PageResultType {
  currentPage* integer
  totalLinks* integer
  totalPages* integer
  result Result {
    title: Result
    default: List []
    ScholixType > {...}
  }
}

```

Figure 4 Response schema for API method /v2/Links

For example, the following snippet shows the response if a client requests the links for the publication “OpenAIRE’s DOIBoost - Boosting CrossRef for Research”, whose DOI is 10.5281/zenodo.1441072:

```
{
  "currentPage": 0,
  "totalLinks": 11,
  "totalPages": 1,
  "result": [
    {
      "RelationshipType": {
        "Name": "IsSupplementedBy",
        "SubType": "isSupplementedBy",
        "SubTypeSchema": "datacite"
      },
      "source": {
        "Type": "literature",
        "Identifier": [
          {
            "ID": "10.5281/zenodo.1441072",
            "IDScheme": "doi",
            "IDURL": "http://dx.doi.org/10.5281/zenodo.1441072"
          }
        ],
        "Title": "OpenAIRE's DOIBoost - Boosting CrossRef for
Research",
        "Creator": [],
        "PublicationDate": "2018-10-01",
        "Publisher": [
          {
            "name": "Zenodo",
            "identifier": []
          }
        ]
      },
      "target": {
        "Type": "dataset",
        "Identifier": [
          {
            "ID": "10.5281/zenodo.1438356",
            "IDScheme": "doi",
            "IDURL": "http://dx.doi.org/10.5281/zenodo.1438356"
          }
        ],
        "Title": "Doiboost Dataset Dump",
        "Creator": [
```

```

    {
      "Name": "Andrea Mannocci",
      "Identifier": []
    },
    {
      "Name": "La Bruzzo, Sandro",
      "Identifier": []
    },
    {
      "Name": "Paolo Manghi",
      "Identifier": []
    },
  ],
  "PublicationDate": "",
  "Publisher": [
    {
      "name": "Zenodo",
      "identifier": []
    }
  ]
},
"HarvestDate": "2019-12-16",
"LicenseURL": null,
"LinkProvider": [
  {
    "name": "Datasets in Datacite",
    "identifier": [
      {
        "ID": "dli_____::datacite",
        "IDScheme": "dnetIdentifier",
        "IDURL": null
      }
    ]
  }
],
"LinkPublicationDate": "2019-12-16"
},
{
  . . .
}
]
}

```

The service has counted close to 2.5Bi hits since January 2018.

## 3.2 Proposed Enhancements

Currently, there are no further improvements to communication and collaboration in the context of OpenAIRE ScholeXplorer.

## 3.3 Integration in production

OpenAIRE ScholeXplorer is already in production (<https://scholexplorer.openaire.eu/>) and supports the Scholix guidelines, for more information please refer to <http://www.scholix.org/guidelines>.

Also, Datacite developed a prototype for exchanging links according to the Scholix guidelines, for more information please refer to <https://www.datacite.org/eventdata.html>.

## 4 | REFERENCES

- [1] Kaschek R., Schewe KD., Thalheim B., Zhang L. (2004) Integrating Context in Modelling for Web Information Systems. In: Bussler C.J., Fensel D., Orłowska M.E., Yang J. (eds) *Web Services, E-Business, and the Semantic Web*. WES 2003. Lecture Notes in Computer Science, vol 3095. Springer, Berlin, Heidelberg, [10.1007/978-3-540-25982-4\\_8](https://doi.org/10.1007/978-3-540-25982-4_8)
- [2] Thalheim B. (2011) The Enhanced Entity-Relationship Model. In: Embley D., Thalheim B. (eds) *Handbook of Conceptual Modeling*. Springer, Berlin, Heidelberg, [10.1007/978-3-642-15865-0\\_6](https://doi.org/10.1007/978-3-642-15865-0_6)

## 5 | FIGURES, TABLES, LISTINGS

- Fig.1 Article on EXPLORE.openaire.eu:  
[https://explore.openaire.eu/search/publication?articleId=od\\_2386::f3a57e5dc273ad9b019d07aaffc2ef7c](https://explore.openaire.eu/search/publication?articleId=od_2386::f3a57e5dc273ad9b019d07aaffc2ef7c)
- Fig.2 Functional abstract view of the article shown at EXPLORE.openaire.eu
- Fig.3 Table of archiving policy state (screenshot)
- Fig.4 Results of the search in a browser
- Fig.5 Results of Restrictions of the search in a browser (partial)
- Fig.6 General conditions of restrictions as results of the search in a browser
- Fig.7 Mockup of scenario 1
- Fig.8 Functional abstract view of scenario 1: simple interlinkage
- Fig.9 Functional abstract view of scenario 2: simple interlink with information snippet
- Fig.10 Functional abstract view of scenario 3: advance interlink with scholarly information snippet
- Fig.11 Functional abstract view of scenario 4: connect to broker service
- Fig.12 Functional abstract view of scenario 5: improve advance scholarly interlinking
- Tab.1 Result of the API query call
- Tab.2 Restriction results of the API query call (partial - pre-/postprints)
- Tab.3 Restriction results of the API query call (partial - conditions)
- List.1 Result of the API query call
- List.2 Restriction result set in the API query call (partial)
- List.3 Condition result set in the API query call (partial)