

Kent Academic Repository

Full text document (pdf)

Citation for published version

Cooper, Helen and Jennings, Liz (2016) ORCID EPrints Implementation Survey Analysis. Working paper. University of Central Lancashire/University of Bath, UK <https://doi.org/10.17030/clok.00015668> <<https://doi.org/10.17030/clok.00015668>>.

DOI

<https://doi.org/10.17030/clok.00015668>

Link to record in KAR

<http://kar.kent.ac.uk/63568/>

Document Version

Publisher pdf

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

ORCID EPrints Implementation Survey Analysis

Background

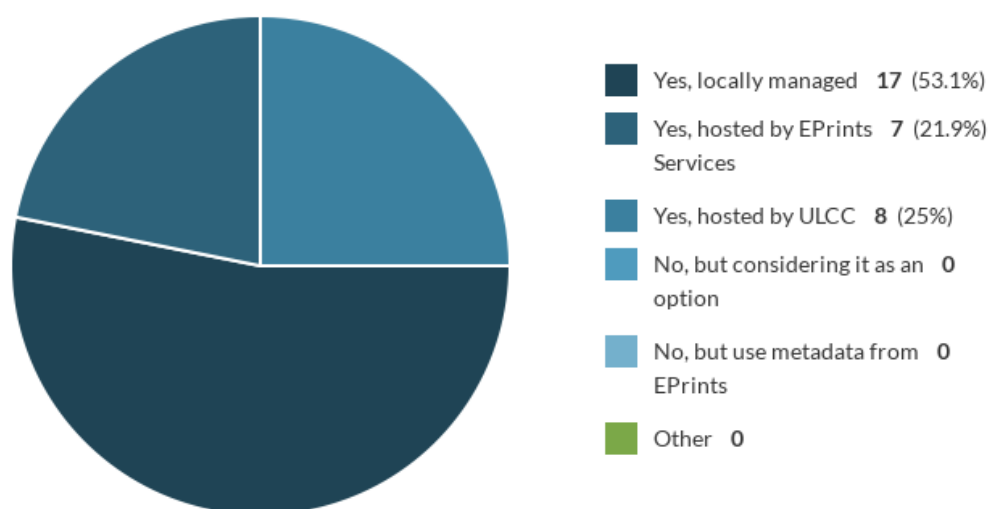
EPrints repository software¹ is used by a wide range of institutions around the world to provide Institutional Repositories, Data Archives, Digital Collections and a range of specialist document stores. The core software is enhanced using plugins, some of which are released to the Bazaar² to be installed in other instances of EPrints. However, most installations have local code which provides for specific requirements that is not easily shared with other institutions.

A number of institutions have already begun the process of integrating ORCID³ functionality into EPrints⁴. Most of the work has been at a local level, although some code and Bazaar packages have already been shared⁵. In order to discover what the community needed from an ORCID EPrints integration we canvassed for use cases⁶ and then applied these responses to a questionnaire, which we made available through EPrints and general repository and Open Access mailing lists⁷.

Current Status

The survey ran from 15th July – 31st August 2016⁸. Thirty responses came from unique organisations, including six from outside the UK. The respondents described themselves as mainly Repository Managers and Librarians or Library Managers, but about a third were developers or technicians. All the institutions have EPrints platforms; a little over half are locally managed and the rest split between the main two hosting organisations ULCC and EPrints Services (Q3).

Q3 Does your organisation have EPrints?



¹ EPrints: <http://www.eprints.org/>

² EPrints Bazaar: <http://bazaar.eprints.org/>

³ ORCID: <http://orcid.org/>

⁴ Wiki page collecting information about ORCID in EPrints: <https://wiki.eprints.org/w/ORCID>

⁵ Import from ORCID (Tier 1 API): <http://bazaar.eprints.org/354/>

ORCID Tier 1 Importer: https://github.com/eprintsug/orcid_tier_1_importer

Login via ORCID: <https://github.com/eprintsug/loginViaOrcid>

ORCID Tier 2: <https://github.com/eprintsug/orcidt2>

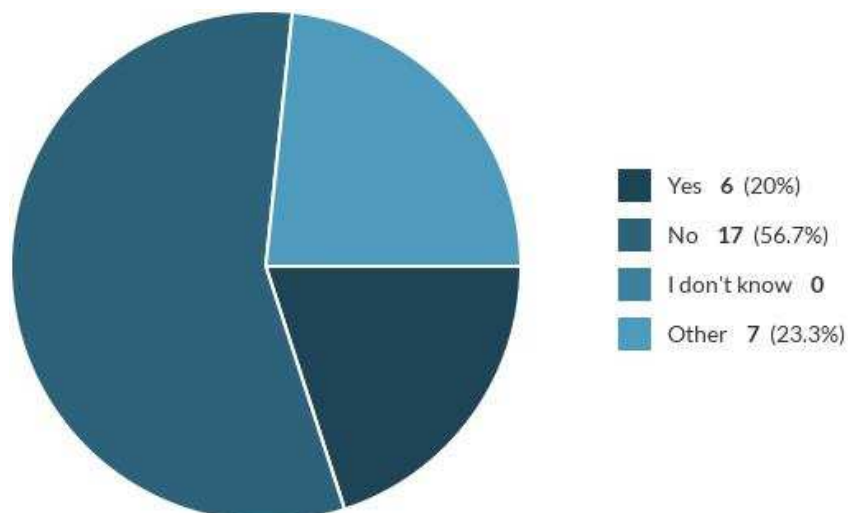
⁶ Wiki page collecting use cases: https://wiki.eprints.org/w/ORCID_connector

⁷ eprints-uk-user-group@googlegroups.com, eprints-tech@ecs.soton.ac.uk, ORCID-UK@JISCMail.AC.UK, UKCORR-DISCUSSION@JISCMail.AC.UK

⁸ Dataset: <http://doi.org/10.15125/BATH-00253>

Over half of the institutions have not yet integrated ORCID with their EPrints repositories and of the rest about half are using very basic functions or are at very early stages of development(Q5).

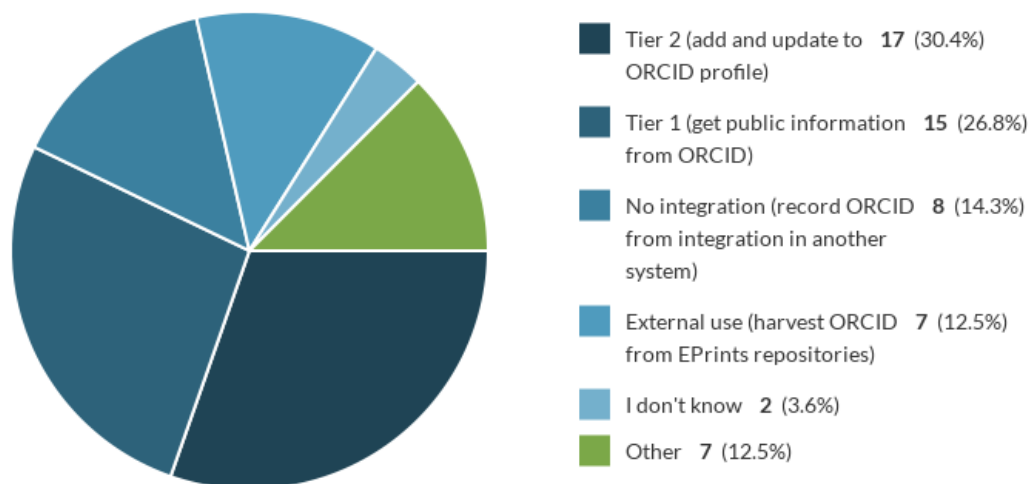
Q5 Have you already implemented an ORCID integration in EPrints?



Use Cases

We asked, “Which types of ORCID Integration are you interested in for EPrints?” (Q4). One third of respondents declared that they want Tier 1 integration (Get public information from ORCID, including identifiers and publications), another third wanted Tier 2 functionality (retrieve ORCID identifiers and synchronise publications and affiliation details).

Q4 Which types of ORCID integration are you interested in for EPrints?



A third group brings together ‘other’ uses including things covered by the APIs, uses of the data once gathered using one or other of the APIs and ‘don’t knows’. Eight users do not want an integration as they have Current Research Information Systems (CRISs) or use other institutional systems as the ORCID end-point (Q4b).

These responses (Q4a) combined with the answers to Q6 (“Please add any further details of your interest in integrating EPrints with ORCID”) can be categorised as three main types of interest in using ORCID in EPrints:

1. Straightforward recording of the identifier in EPrints and using the identifier as a way of disambiguating authors. An attractive way to display the identifier in exports, feeds like MePrints profile pages, and reports was also desirable.

2. More advanced exchange of information between EPrints instances and ORCID and a way for researchers to register with ORCID via EPrints
3. Developers looking to share existing work and looking for user feedback – also users looking for more documentation accompanying the things that already exist. Bazaar packages and GitHub code are not in a finished enough state for the non-developer to use and developers have limited time to take the work they do beyond a non-local stage to general user friendliness.

These main types can be loosely described as:

1. Tier 1 functionality
2. Tier 2 functionality
3. The technical solution

What they all have in common is a need for consistency, user friendliness, documentation and dissemination.

Specific Features Required

Questions 7 and 8 asked respondents about the relative importance of various features harvested from the use cases document⁹ relating to how ORCID identifiers are recorded and how they are displayed.

Essential or important features are:

- Record ORCID identifiers as an additional subfield against the creator (Pittsburgh approach¹⁰) *
- Record field includes validation of ORCID identifiers¹¹
- Record field includes ORCID identifier look-up (prioritises local authors first)
- Display ORCID identifier in EPrint abstract page
- Display ORCID identifier in page metadata
- Render ORCID as a link to ORCID profile*
- Include ORCID in all metadata export formats (JSON, CSV)*

*No-one identified these features as unimportant

Features that are just useful are:

- Record ORCID identifier in a new field
- Display ORCID in user profile for visitors
- Display ORCID identifier in citation

Not important features:

- Recording the identifier using the email subfield

They were also asked about the features present in the Tier 1 API and the Tier 2 API.

Tier 1: Most people wanted to use Tier 1 or they weren't sure. Those that were definite negative or unsure were mainly intending to use another system or were looking to use Tier 2. A relatively large proportion didn't know.

⁹ Wiki page collecting use cases: https://wiki.eprints.org/w/ORCID_connector

¹⁰ Pittsburgh implementation notes: https://wiki.eprints.org/w/ORCID#Notes_from_Pittsburgh

¹¹ Structure of an ORCID identifier: <http://support.orcid.org/knowledgebase/articles/116780-structure-of-the-orcid-identifier>

It was mainly considered essential or important that administrators should be able to import items from ORCID but even distributed between essential/important and useful/not important that depositors (*i.e. authors?*) should be able to import items from ORCID.

Tier 2: An even larger proportion stated that they didn't know if they wanted to use the Tier 2 API – over half. However, again most felt that the features in Tier 2 were important to have – namely the ability for administrators to export items to ORCID and the ability to record permissions relating to ORCID accounts in EPrints.

Additional requirements not explicitly asked about are:

- Automatic import/export and alignment – non-duplication
- Any solution needs to be consistent across the community and easily upgraded or included in regular upgrade packages provided by the hosting services: *i.e.* needs to be supported and adopted by hosting service providers).
- Reporting ORCID data to IRUS
- Disambiguation

Conclusion

Functionality

From this survey it seems that there is little variation in what most people want from ORCID/EPrints integration. The basic ORCID functions of offering disambiguation and identity authorisation are very desirable. Integration needs to provide:

1. Attractive and useful recording and display of accurate ORCID identifiers that allow users to be clear about the identity of authors in the same way that we can record, display and use other unique identifiers like DOIs. This includes a seamless way for authors to register/align their EPrints records with ORCID identifiers.
2. Smooth administrative functions, including downloading information from ORCID and other sources, reporting and de-duplication. Good, clear documentation for administrators is especially important.

These two areas roughly map to the existing tier 1 and tier 2 APIs but this terminology does not appear to be very useful as the functionality would be more desirable in a simple, seamless offering. Many people who have adopted the APIs seem to be waiting on further development – wanting a more finished product than is currently available – which suggests that the current offering requires too much development to make it work. This is also leading to variety and less than optimum use of resources.

Development

There appears to be a lot of development that has happened in various institutions. However, this is aimed at tight local specifications and addresses different areas, or has had success in different areas depending on the skills/interests of the developer. There does not seem to be a need for new code, but for a project to bring all the existing code together to create a consistent and shareable finished product that takes advantage of all the best aspects of existing work. All the developers are keen for no new-wheel-inventing to be funded, but for their own work to be carried forward and melded with that of others.

Recommendations

Based on our research, we recommend that JISC should act as a co-ordinator to harness all the development work into a single project. There are two main strands of work that should be prioritised:

- 1) Produce a single product that combines existing work into a single Bazaar package to answer the needs of the community as detailed above
- 2) Fund comprehensive documentation to help us all make the best use of the work done. The JISC consortium website is the perfect place for this.

26th September 2016

Helen Cooper
Repository Manager & University Archivist
University of Central Lancashire
hcooper@uclan.ac.uk
<http://orcid.org/0000-0002-2968-4747>

Lizz Jennings
Technical Data Officer
University of Bath
e.jennings@bath.ac.uk
<http://orcid.org/0000-0002-4052-3602>

Licence

This report is made available under a [Creative Commons Attribution 4.0 Licence \(CC-BY 4.0\)](#).

Data Availability

Data collected in the survey are available from <http://doi.org/10.15125/BATH-00253>