

Grant agreement no. 312788



**ORCID AND DATACITE
INTEROPERABILITY NETWORK**

www.odin-project.eu

D1.2 First interim scientific report

WP1 – Project Management

V2_0

Final

Abstract:

This first interim scientific report, as a contractual deliverable, describes at a high level the work that the ODIN consortium has carried out in the first 6 months of the project, and an overview of the resources used in the project so far.



Lead beneficiary: DataCite

Date: 11/03/2013

Nature: Report

Dissemination level: PU (Public)

ODIN is co-funded by the EC under the e-Infrastructures Activity of the FP7 Capacities Specific Programme

© 2013 ODIN Consortium. Some rights reserved. This work is licensed to the public under the Creative Commons Attribution 3.0 License. <http://creativecommons.org/licenses/by/3.0/>



	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final 2/24

Document Information

Grant Agreement no.	312788	Acronym	ODIN
Full title	ORCID and DataCite Interoperability Network		
Project URL	http://odin-project.eu		
Project Coordinator	John Kaye (BL) Address: The British Library 96 Euston Road, London NW1 2DB, United Kingdom Phone: +44 20 7412 7450 Email: john.kaye@bl.uk		

Deliverable	Number	1.2	Title	First interim scientific report
Work package	Number	1	Title	Project Management

Document identifier	ODIN-WP1-DEL-0002-0_1			
Delivery date	Contractual	Feb 2013	Actual	04 March 2013
Status	Version 1.0		Final <input checked="" type="checkbox"/> Draft <input type="checkbox"/>	
Nature	Report <input checked="" type="checkbox"/> Prototype <input type="checkbox"/> Demonstrator <input type="checkbox"/> Other <input type="checkbox"/>			
Dissemination Level	<input checked="" type="checkbox"/> Public <input type="checkbox"/> Restricted to other programme participants (including the Commission Services) <input type="checkbox"/> Restricted to a specified group (including the Commission Services) <input type="checkbox"/> Confidential, only for consortium members (including the Commission Services)			

Authors (Partner)	Sergio Ruiz (DataCite), John Kaye (British Library)			
Responsible Author	John Kaye		Email	John.kaye@bl.uk
	Partner	The British Library	Phone	+44 7412 7450

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

Document Status Sheet

Issue	Date	Comment	Author
0_1	06 Mar 2013	Initial draft	Sergio Ruiz (DataCite)
0_2	07 Mar 2013	Revised Draft with input from WP's	John Kaye (BL) Laura Rueda (CERN) Suenje Dallmeier-Tiessen (CERN) Gudmundur Thorisson (ORCID)
0_3	07 Mar 2013	Revised Draft with CERN review	Laura Rueda (CERN) Suenje Dallmeier-Tiessen (CERN)
0_4	08 Mar 2013	Revised Draft with ORCID review	Gudmundur Thorisson (ORCID)
0_5	08 Mar 2013	Revised Draft with WP4 and open consultation input and BL review	John Kaye (BL) Suenje Dallmeier-Tiessen (CERN) Martin Fenner (ORCID) Laure Haak (ORCID) Jan Brase (DataCite)
1_0	08 Mar 2013	Final Version	Sergio Ruiz (DataCite)
2_0	11 Mar 2013	Final Version	Salvatore Mele (CERN) Martin Fenner (ORCID) John Kaye (BL)

Document Change Record

Issue	Item	Reason for Change
2_0	Whole document	Updated details on use of resources and additional cosmetic changes

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

TABLE OF CONTENTS

1. SUMMARY	5
2. PROJECT OBJECTIVES, WORK PROGRESS AND ACHIEVEMENTS	8
2.1. PROJECT OBJECTIVES FOR THE PERIOD	8
2.1.1. Access	8
2.1.2. Discovery	9
2.1.3. Interoperability	9
2.1.4. Sustainability.....	10
2.2. WORK PROGRESS AND ACHIEVEMENTS DURING THE PERIOD	11
2.2.1. WP2 – Communication	11
2.2.2. WP3 - Proofs of Concept	13
2.2.3. WP4 – Interoperability.....	16
2.2.4. WP5 – Strategy	18
2.2.5. WP6 – Internationalisation	19
2.3. PROJECT MANAGEMENT DURING THE PERIOD	21
2.4. DELIVERABLES AND MILESTONES TABLES.....	23
2.4.1. Deliverables	23
2.4.2. Milestones.....	24

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

1. SUMMARY

ODIN: ORCID and DataCite Interoperability Network – <http://www.odin-project.eu>

Context

‘Data as infrastructure’ is a critical concept for a fully-integrated European Research Area (ERA) to drive innovation forward as envisaged by the Digital Agenda for Europe. The lack of data availability hinders this vision. In academic publishing, peer review and citation have long been recognised as mechanisms for endorsing the trustworthiness of research outputs and incentivizing researchers to contribute. Trustworthy research data will only be widely available if the same principles are applied.

Key, participative, initiatives have emerged to address this challenge. The DataCite consortium has assigned over one million DOI names in the last few years to make research data citable. The Open Researcher and Contributor ID (ORCID) was launched in October 2012 and by the end of February 2013 has over 75,000 researchers and contributors registered. ORCID offers the opportunity to identify individual authors across systems and infrastructures, including scholarly works that can have up to thousands of authors (as in the case of the High-Energy Physics community). Researchers often change their affiliation, and collaborate across national disciplinary boundaries.

The ORCID and DataCite Interoperability Network (ODIN), a project funded by the European Commission (FP7), aims to build on the success of DataCite and ORCID by designing an ‘awareness layer’ for persistent author and object identifiers, thereby reducing technical, cultural and logistical barriers to the accessibility, attribution and trust of data. Identifier awareness will make it possible to stabilise: References to a data object; Tracking of use and re-use; Links between a data object, subsets, articles, rights statements and every person involved in its life-cycle (creator, editor, reviewer, aggregator, etc.).

Given the importance of these functions as we approach HORIZON2020, we aim to prove the feasibility of author, data and rights identification, promote trust building towards open scientific data e-Infrastructures and lay the foundation necessary to promote future interoperability (technical, semantic, reference architecture, etc.) in the scientific data domain in Europe and globally.

The project is coordinated by The British Library and the ODIN consortium is formed by The British Library, CERN, DataCite, ORCID EU, The Australian National Data Service, arXiv and Dryad.

Work performed during the first 6 months

The project got off to a good start with the planning and holding of a successful kick-off meeting that provided an overview of the current persistent identifier landscape, as well as valuable input from key stakeholders to the project.

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

A communications plan and website were produced for the purpose of effectively communicating ODIN's findings and events to the relevant audiences and communities throughout the project. Work has begun to develop two disciplinary proofs of the concept of open and interoperable persistent identifiers of data and contributors in scholarly communication. The planned work on the interoperability framework between open identifiers for data and contributors in different infrastructures is well under way. The detailed assessment of the gaps that hinder the implementation of the "thin layer" of interoperable, open, persistent identifiers started with the input of the kick off meeting. In parallel the roadmap is being produced, which will include recommendations for addressing and closing of such gaps, and the creation of a network of expertise and connections to relevant stakeholders.

Main results

Three deliverables were submitted to the European Commission during the first six months, as stipulated in the Description of Work. These are:

D1.1 Project Quality Plan & Project Risk Register. Defines the project management structure, procedures, organisation and the methodology that all the partners shall apply throughout the project.

D2.1 Kick-off preparation, Communication plan and Website. Describes ODIN's approach to engaging with and disseminating project outputs to key target audiences via various channels for the two-year duration of the project. It is important to remark that the first two crucial items in this plan have been completed as expected: holding the first of three major community events (the "kick off" meeting) and launching a project website (<http://www.odin-project.eu>).

D2.2 Kick-off Report. The ODIN kick-off meeting provided an overview of the current persistent identifier discussion, as well as valuable input from key stakeholders to the ODIN project. Eleven presenters from a variety of stakeholder groups provided important perspectives, and the extensive discussion by all meeting participants further elaborated on key points. Based on the key scientific inputs, the following actions were identified:

- i) Encourage discussion and collaboration between communities
- ii) Address specific issues of working with large numbers of small heterogeneous datasets
- iii) Identify main interoperability use cases and define their business value
- iv) Encourage development of novel tools to track the reuse of data

All public deliverables can be found at <http://odin-project.eu/project-outputs/>

Future Work

Work is in progress on the two disciplinary proofs of concept, the conceptual model of interoperability and the gap analysis and draft roadmap, all due in Month 11 (August 2013). In this context, and as stipulated in the DoW, representatives from non-European

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

partners will visit their European counterparts in Spring 2013 to provide international input into the work carried out in each of the areas.

An integral part of this work will be an open consultation process which will be held for the gap analysis and draft roadmap and the conceptual model of interoperability. This will include an open online discussion with the community. In addition, views and input from selected experts will be sought on the occasion of the ORCID Outreach Meeting, Symposium on Research Attribution, and the ORCID Codefest, in May 2013 in Oxford, UK. While not organised within the ODIN project, these events present a unique opportunity to access the leading experts in this field, to keep a continuous relation with the community of the ODIN stakeholders.

Dates for the ODIN “Big Bang” 1st year conference and codesprint (hackathon), a key deliverable from the project first year, have been set for 15th – 17th October 2013. This and other event information has been published on the project website at <http://odin-project.eu/events/1st-year-big-bang-and-codesprint/>.

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

2. PROJECT OBJECTIVES, WORK PROGRESS AND ACHIEVEMENTS

2.1. Project objectives for the period

The ODIN Project identified four specific objectives to accomplish at the end of the project. These are:

- Access
- Discovery
- Interoperability
- Sustainability

We present details for each of those objectives and the corresponding achievements to date.

2.1.1. Access

Researchers must be able to seamlessly access datasets used in a journal article, a grant report, or another scholarly artefact, with explicit and unambiguous terms of use.

ODIN objective: Build shared understanding

The project will engage with all ODIN stakeholder communities to understand how gaps in the support for data and contributor identifiers may be bridged, in order to meet the needs of the community for access to research data. Consensus will be built on a road-map to implement the missing identifier layer, and overcome the access challenge. This will be subject to a thorough validation, and fully-supported by all stakeholders, with concrete recommendations feeding into HORIZON2020.

ODIN held a successful kick-off event where key stakeholders in the community contributed their thoughts and ideas. This validated the project's aims and objectives and led to consensus about what needs to be achieved.. This engagement has continued through communications activities by each partner. The forthcoming consultation process for WP4 and WP5 will ensure further shared and feedback. This open consultation process comprises the following steps:

Based on the ODIN research, first drafts of the respective deliverables are developed. These are first discussed internally, with input from the international partners (WP6). These drafts are then shared to be commented on and discussed by the relevant communities. To ensure the participation of relevant experts and key stakeholders ODIN project partners will invite such groups specifically. This approach ensures that ODIN views are broad enough to cover enough disciplines and detailed enough to understand the current and forthcoming challenges.

The draft will be revised by ODIN project members in an iterative process; comments and suggestions will be incorporated.

For WP5, a second step is envisioned through direct interviews with experts already gathered for a relevant meeting, not organised by ODIN: the ORCID Outreach Meeting,

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

Symposium on Research Attribution, and the ORCID Codefest, in May 2013 in Oxford, UK. This step informs in particular the roadmap that is being drafted.

This consultation will avoid silo developments and foster a common vision that has already informed and will continue to inform the Horizon 2020 strategy. The consultation process will also further collaborations with relevant partner projects and stakeholder so that a broader, but detailed, input to H2020 can be given.

2.1.2. Discovery

Researchers must be able to identify scholarly works or datasets related to each other either through networks of references and citation or through shared contributors

ODIN objective: explore proofs of concept

The project will analyse two very different disciplines

- *Humanities and Social Sciences (HSS)*, where generators, curators and users of multiple datasets are separated by considerable distances in space and/or in time, without any prevailing framework for identification of the data or the different contributors.
- *High Energy Physics (HEP)*, where massively multi-authored publications and their accompanying datasets exist in multiple versions across several repositories, with no mechanism currently in place to aggregate them.

The ability to discover, navigate and represent the association between datasets and contributors will be achieved for the first time in these fields. The lessons learnt, and commonalities, will feed into the ways in which we address the other challenges

The main goal of this is to develop disciplinary proofs of concept for open and interoperable persistent identifiers of data and contributors in scholarly communication. In particular, proving the ability to navigate across data and contributors in the area of High-Energy Physics (HEP) and Humanities and Social Sciences (HSS) and exploring commonalities with each other and other knowledge areas.

The process of designing the models and building the proof of concept started after the kick-off meeting and is now well underway, thanks to intense contacts of each partner with all stakeholders in the area..

2.1.3. Interoperability

Users must be able to connect datasets and contributors across independent platforms using different, but interoperable identifier schemes.

ODIN objective: demonstrate interoperability

DataCite and ORCID are participative infrastructures where each member organisation has full control of its identifier system. DataCite already allows datasets to be identified

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

across systems, while the ORCID design allows profiles to be built and matched to individual identifiers.

As a first step, ODIN partners have identified different use cases of interoperability between data and author identifiers. These use cases are the foundation of the ongoing work on the conceptual model of interoperability. Designing concrete interoperability solutions between ORCID, DataCite and other identifiers, so that data in selected systems can be connected to contributors identified other systems, and vice versa.

2.1.4. Sustainability

Funding agencies and research institutions must be able to depend on the on-going viability of the systems that provide and maintain persistent identifiers. The organisations that deliver those services may be commercial or non-profit, but in either case their sustainability depends on the value of the services they provide.

ODIN objective: stimulate added-value services

There are enormous opportunities for the providers of identifiers, and for innovative third parties, to provide value-added services on top of a layer of interoperable data and contributor identifiers.

ODIN has already engaged with existing providers of products and services within the e-Science domain, ranging in scale from individual developers to large corporations. Starting with a 'big bang' the ODIN Codesprint (Hackathon), ODIN will explore potential scenarios for exploitation of unique attribution, such as evaluation of the impact of datasets associated with specific contributors. Concrete examples will be developed and demonstrated during the project lifetime. The participation of several ODIN partners to the ORCID Outreach Meeting, Symposium on Research Attribution, and the ORCID Codefest, in May 2013 in Oxford UK will allow to identify potential candidates for the participation to the 'big bang' ODIN Codesprint (Hackathon).

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

2.2. Work progress and achievements during the period

2.2.1. WP2 – Communication

Summary of progress towards objectives and details for each task/significant results

The main objective of WP2, as per the DoW, is to offer a communication framework within and beyond the ODIN Consortium in order to ensure representation of all stakeholder groups in the activities of the project.

During the first sixth months of the project we have communicated to stakeholder communities various information on published project outputs, upcoming events and other activity. Specifically, we have done the following as per our Communication plan (see D2.1):

- launched a project website at <http://odin-project.eu>
- published deliverable D2.2 (kick-off report) via the Figshare online repository (see <http://dx.doi.org/10.6084/m9.figshare.107019>)
- published several blog posts on the website to announce availability of D2.1, upcoming ODIN events and more
- communicated the above to target audiences as appropriate (E-mail lists, Twitter etc.) via individual ODIN partners

Activities by task

Task 2.1: Preparation and management of kick-off event, M1 – M2 (ORCID EU, DataCite, all)

The kick-off meeting was held as scheduled and provided an overview of the current persistent identifier discussion, as well as valuable input from key stakeholders to the ODIN project. Eleven presenters from a variety of stakeholder groups provided important perspectives, and the extensive discussion by all meeting participants further elaborated on key points. Based on the key scientific inputs, the following actions were identified:

1. Encourage discussion and collaboration between communities
2. Address specific issues of working with large numbers of small heterogeneous datasets
3. Identify main interoperability use cases and define their business value
4. Encourage development of novel tools to track the reuse of data

These actions and scientific inputs from speakers and participants at the kick-off event have fed into the thinking in each of the work packages.

A full report was provided in the form of deliverable D2.2 (Kick-off Report). The report was sent to the EC, published online (<http://dx.doi.org/10.6084/m9.figshare.107019>) and disseminated as noted above.

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

Task 2.2: Communication and dissemination, M1 – M23 (ORCID EU, DataCite, all)

The Communication plan was put into place. The project website was also set up at (www.odin-project.eu) and is being updated on a regular basis. ODIN partners are constantly identifying and communicating with the different communities of interest for the project. Deliverable D2.1 (Kick off preparation, Communication plan and Website) was produced and sent to the EC.

Task 2.3: Planning and management of “big bang”, M3 – M11 (ORCID EU, CERN, all)

The first year conference was initially scheduled for month 11 (August 2013), but this has been rescheduled for month 14 (15-18 October 2013). These dates were agreed with the Project Officer. CERN, in cooperation with ORCID EU, is already working on a list of potential participants to invite, potential key areas to work on and other preparation work. This event is being organised bearing in mind the proofs of concept from WP3, as well as the preliminary gap analysis and roadmap already being developed by WP5.

Dates for the event have been announced via the event page on the website (see <http://odin-project.eu/events/1st-year-big-bang-and-codesprint/>) and a blog post, and also communicated more widely by partners as noted above. Additional event promotion will take place as we draw closer to the event and (for example) agenda information is available for communicating to potential participants. The participation of several ODIN partners to the ORCID Outreach Meeting, Symposium on Research Attribution, and the ORCID Codefest, in May 2013 in Oxford UK will allow to identify potential candidates for the participation to the ‘big bang’ ODIN Codesprint (Hackathon).

Deviations from DoW/ impact on other tasks

D2.1 and D2.2 were delivered with marginal delay due to longer-than-expected lead time to hire ORCID EU staff, who joined in October. This has had no impact on other contracted deliverables.

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

Use of resources (including actual vs. planned)

Table of person-months per partner compared with the DoW.

No.	Short name	PM expected to be spent during the first 6 months of the project	PM spent in WP2
1	ANDS	0.0	0.0
2	arXiv	0.0	0.0
3	BL	0.3	1.1
4	CERN	0.8	0.2
5	DataCite	1.0	0.5
6	DRYAD	0.0	0.0
7	ORCID	3.3	4.2
TOTAL		5.3	6.0

2.2.2. WP3 - Proofs of Concept

Summary of progress towards objectives and details for each task/significant results

The main objective of WP3, as per Annex 1 is to develop two disciplinary proofs-of-concept for open and interoperable persistent identifiers of data and contributors in scholarly communication, in a variety of current and future scenarios.

After input from the kick-off meeting, work has progressed toward building the proofs of concept in both HSS and HEP.

This work package has no deliverables in this period, but the process of designing the models and building the proof of concept started after the kick-off meeting.

Activities by task

Task 3.1 Common kick-off, M1 – M2 (BL, CERN, DataCite)

In order to obtain the widest perspectives, external stakeholders were invited to the kick-off meeting. Regarding the WP3 point of view, their input granted the extrapolation of the proofs of concepts to more areas than HSS and HEP. Service providers and publishers exposed their current status and their next steps towards the interoperability of systems; researchers, librarians and data centres stressed their needs for the future of data sharing; and funders provided a community-focused context and encouraged the use of permanent identifiers.

For the HSS proof of concept, the meeting provided many useful insights from a wide range of stakeholders. Of particular relevance was the input of GESIS, SSRN, JISC and UK REPOSITORYNET+. GESIS explained how they use DataCite DOIs for datasets in German economic and social collections, with many commonalities shared with the approach by ESDS in the UK. SSRN highlighted the need to show the benefits of sharing research data to the social science community and gave useful inputs from a

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

US perspective. JISC's input was very useful for this work package as it allowed us to see their current thinking around name and object identifiers which will help us to make choices in the development of the conceptual model and inform workflows from a UK perspective.

JISC also expressed interest in engaging with the ODIN project because working with names and objects bridges 2 different areas of work for them (name identifiers and research data management). It was also useful to hear about the work of UK REPOSITORYNET+ as many of the objects relevant to the HSS proof of concept are held in UK repositories.

For the HEP proof of concept, the kick-off meeting provided invaluable input from different stakeholders and scenarios, allowing the opportunity to compare issues, needs and workflows. This cross-disciplinary view applies directly to the design of the proof of concept, where the distinctive features of HEP have clear similarities with other areas, such as dataset versioning in the social sciences, increasing amounts of data sent to publishers and difficulties in disambiguating authors, count citations, etc. This situation offers an opportunity to lead the way towards the adaptation and unification of workflows, as is a necessary basis for the clearly demanded interoperability of systems.

Task 3.2 HSS. M2 – M23 (BL), First phase M2-M11

The first phase of Task 3.3 lasts until month 11. This task looks at creating a conceptual model around the use and re-use of British Birth Cohort Studies datasets. In the first 6 months, workflows for the conceptual model using ORCID and DataCite have been designed and a number of challenges have been identified.

Datasets in the proof of concept are affected by different policies; for example, surveys hosted by the Economic and Social Data Service (ESDS) have DataCite DOIs and open metadata, whereas surveys funded by the Medical Research Council (MRC) do not use DataCite DOIs and access to metadata requires approval. Through ODIN we have opened up a dialogue with the MRC surveys to see how they can be integrated into the conceptual model. We have found that quality of metadata concerning data creation and curation roles is variable and we are exploring ways to improve this. Due to the long history of this dataset, data creators, contributors and authors may not be currently engaged in research. Therefore ORCID identifiers for those individuals may not be a feasible option and we will thus be bringing ISNI identifiers, JISC Names or other person identifiers into the model.

Through data collection for the model, we have successfully made links between surveys and datasets and different types of publication (published journals, working papers, grey literature etc.), which is essential to track the impact of data resources. These outputs use different identifiers, which will be reflected in the model.

In M9 a representative from the Australian Data Archive will be visiting BL for three weeks. During this time they will provide input into the conceptual model and meet and share experiences with data creation and curation partners involved in the proof of concept. Also in M9, ODIN will be running a session on identifiers at the International

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

Association for Social Science Information Service and Technology (IASSIST) conference.

Task 3.3 HEP M3 – M23 (CERN, DataCite), First phase M2-M11

Conceptual models for the particularities of the High-Energy Physics fields have been designed and discussions have already taken place with the relevant stakeholder and Cornell-arXiv, the non-European partner.

Scholarly communication in HEP needs to deal with the “hyperauthorship” phenomenon (several thousand authors per article is not uncommon) and the extension of this to the datasets generated by the multitude of collaborations running the experiments. This data, with different levels of abstraction, lives in separate, large-scale data centres and the community involved in the generation of these datasets sees the implementation of permanent identifiers for both data and authors/creators as a big opportunity to facilitate access.

In this situation, every datacentre uses different, sometimes proprietary and in general non-interoperable identifiers. The potential interrelation between them needs to be based in common identification systems. The first conceptual models for them will be based on identifiers from DataCite and ORCID.

Deviations from DoW/ impact on other tasks

The BL has dedicated more effort than initially planned in this initial phase of the project, to fully accommodate the interesting feedback from the kick-off meeting and to establish the links with all stakeholders in the field, to lead to a successful result.

CERN has allocated its own additional resources beyond the project funding envelope to the proof-of-concept, with a junior software engineer working at the interface between the design of the proof of concept and the INSPIRE digital library platform.

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

Use of resources (including actual vs. planned)

Table of person-months per partner compared with the DoW.

No.	Short name	PM expected to be spent during the first 6 months of the project	PM spent in WP3
1	ANDS	0.0	0.0
2	arXiv	0.0	0.0
3	BL	2.0	3.5
4	CERN	2.0	9.1
5	DataCite	1.0	0.1
6	DRYAD	0.0	0.0
7	ORCID	0.0	0.0
TOTAL		5.0	12.7

2.2.3. WP4 – Interoperability

Summary of progress towards objectives and details for each task/significant results

The main objective of this WP is to prove the value of achieving interoperability between open identifiers for data and contributors in different infrastructures.

ORCID and DataCite have been working during the first six months of the project in trying to examine the current state-of-the-art in identifier systems and discuss with different stakeholders the needs and challenges for identifier interoperability.

Activities by task

Task 4.1 Preparation for Kick-Off, M1 – M2 (DataCite, ORCID EU)

The scope of WP4 (Interoperability) is to prove the value of achieving interoperability between open identifiers for data and contributors in different infrastructures. At day one of the kick-off event we received especially valuable inputs on this issue from two external participants: OKKAM which represented the EC-funded DIGOIDUNA study on the role of identifiers for digital objects and authors; and OCLC which represented the International Standard Name Identifier (ISNI) initiative. These inputs will be included in the first WP4 deliverable "Conceptual model of interoperability" due August 2013.

Task 4.2 Conceptual Model, M3 – M8 (DataCite, ORCID EU, all)

In joint meetings, DataCite and ORCID identified at least 3 different use cases of interoperability between data and author identifiers. These use cases are the foundation of the on-going work on deliverable D4.1, describing the conceptual model of interoperability. The conceptual model will define how these identifiers can successfully and openly interoperate inside each of the individual targeted infrastructures and

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

suggest open avenues for their mutual integration. This deliverable will be subject to an open consultation process along with D5.1

DataCite organised a meeting with representatives from the EPIC consortium that offer a handle-based data identification infrastructure (<http://www.pidconsortium.eu/>). Based on this meeting the developers of DataCite and EPIC have discussed the technical concepts of identifier interoperability and methods of mapping one identifier to the other. Use-cases were defined to describe how different identifiers can be used for data in different parts of the research lifecycle and how they can evolve into each other. ORCID EU has worked with ISNI (<http://www.isni.org>) with the aim for interoperation through linking and sharing of public data between the two systems. In addition, ORCID and ISNI are committed to investigate the feasibility of a shared identifier scheme for a single number to represent an individual in both the ORCID and ISNI databases.

During M9 a selected expert from Dryad will be visiting ORCID and DataCite to provide technical input into describing the conceptual model of interoperability.

Deviations from DoW/ impact on other tasks

The recruitment for the project in ORCID, the BL and DataCite took longer than expected. As a result, the existing staff was spread across other work-packages and in particular WP2 (ORCID EU), WP1 and WP3 (BL and DataCite). This resulted less effort than initially expected has been invested in this Work Package. All personnel is by now in place and work is accelerating. This deviation has not further impact in other Work Packages.

Use of resources (including actual vs. planned)

Table of person-months per partner compared with the DoW.

No.	Short name	PM expected to be spent during the first 6 months of the project	PM spent in WP4
1	ANDS	0.0	0.0
2	arXiv	0.0	0.0
3	BL	0.8	0.1
4	CERN	0.3	0.0
5	DataCite	3.0	2.2
6	DRYAD	0.0	0.0
7	ORCID	2.3	0.7
TOTAL		6.3	3.0

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

2.2.4. WP5 – Strategy

Summary of progress towards objectives and details for each task/significant results

The main objective of this work package is to enable the implementation of the ‘missing thin layer’ of interoperable, open, persistent identifiers in the architecture of the Collaborative Data Infrastructure proposed by HLEG in 2010.

The WP5 work is focused on identifying the gaps that hinder implementation of the "thin layer". The idea is to, as a first step, establish a detailed assessment of the current gaps and practices. This will be combined with a roadmap that includes recommendations to address and close such gaps.

During the first 6 months of the project the main objective of WP5 has been to develop a framework to for the gap analysis and the roadmap, to build a network of expertise and connections to relevant stakeholders. The output from the kick-off event framed a good starting point and will have an impact on all first year activities in this WP. WP5 has no deliverables in this period, but work started in a close collaboration with other work packages after the kick-off meeting.

Activities by task

Task 5.1 – Prepare for Kick-Off, M1 – M2 (CERN, all)

The Kick-Off event has been prepared in strong collaboration with WP2. A comprehensive set of speakers has been selected to represent all the relevant stakeholders. The topics of discussion were carefully planned with these experts in advance to the meeting. Speakers with interesting use-cases for persistent identification in scholarly communication contributed to the discussion.

Task 5.2 – Gap analysis and draft roadmap, M2 – M11 (CERN, BL)

The presentations and the discussion from the kick-off event have been used as a first input for the gap analysis. In addition, relevant developments and publications have been incorporated. For example, preliminary results highlight the need for strong collaboration among stakeholders to address the tracking of research materials across platforms and disciplines. Important aspects to consider for the forthcoming work are interoperability and standards, which build a base for cross-disciplinary value added services. These need to connect an open sharing environment and the academic reward and incentive mechanisms.

An open consultation process, together with WP4, has been planned. An important step in this task will be the participation of several ODIN partners to the ORCID Outreach Meeting, Symposium on Research Attribution, and the ORCID Codefest, in May 2013 in Oxford UK, where key international stakeholders are expected to participate. They will be interviewed to gather their views on a first draft of the gap analysis and roadmap.

Deviations from DoW/ impact on other tasks

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

Minor deviations for this WP are an artefact of a linear interpolation of effort over the first year. It is expected that partners beyond CERN contribute expertise and support over the consultation process, to happen later in the first year of the project..

Use of resources (including actual vs. planned)

Table of person-months per partner compared with the DoW.

No.	Short name	PM expected to be spent during the first 6 months of the project	PM spent in WP5
1	ANDS	0.0	0.0
2	arXiv	0.0	0.0
3	BL	0.5	0.0
4	CERN	2.8	2.8
5	DataCite	0.3	0.0
6	DRYAD	0.0	0.0
7	ORCID	0.3	0.0
TOTAL		3.8	2.8

2.2.5. WP6 – Internationalisation

Summary of progress towards objectives and details for each task/significant results

To be successful globally, the solutions proposed by ODIN must address the concerns of those global audiences. WP6 provides an additional mechanism for validating that relevance.

Activities by task

Task 6.1 Preparation for kick off and communication, M1- M2 (ANDS, all)

The Internationalisation work package (WP6) received valuable input at the kick-off meeting from a broader range of international stakeholders beyond those that are formally partners the ODIN project. Global information services such as Elsevier and VIAF/ISNI provided significant perspectives on the scientific challenges of international researcher identification. Differing cultural perspectives were evident in inputs from the USA (SSRN Publishing) and Japan (The National Institute for Informatics) with important commonalities noted in the approaches to tracking datasets across national boundaries.

The second, internal day of the kick-off meeting established reliable intra-project processes for incorporating kick-off inputs and future feedback into the project deliverables (particularly at the critical project junctures of month 12 and 18). Timetables and responsibilities were confirmed for ANDS, ARXIV and DRYAD to both



	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

have input into and review project deliverables with processes for validating proposals from international perspectives.

Task 6.2 Agile validation in the first year, M2-M11 (ANDS, all)

WP6 partners are in regular contact with the rest of the consortium via electronic communications and fortnightly project team conference calls. Visits by WP6 participants to European partners to provide input into work packages have also been arranged. Simeon Warner from arXiv is visiting CERN in March (M7). Steve McEachern from The Australian Data Archive is visiting BL in May (M8). Dryad software developer (Ryan Scherle) is scheduled to meet with ORCID and DataCite staff in the UK for one week in April (M8). Todd Vision is then to visit DataCite, ORCID and CERN over M8 and M9.

Deviations from DoW/ impact on other tasks

No particular deviations have been observed for this WP, as most effort will be deployed during the incoming visits.

Use of resources (including actual vs. planned)

Table of person-months per partner compared with the DoW.

No.	Short name	PM expected to be spent during the first 6 months of the project	PM spent in WP6
1	ANDS	1.0	0.7
2	arXiv	0.5	0.4
3	BL	0.0	0.0
4	CERN	0.0	0.0
5	DataCite	0.0	0.0
6	DRYAD	0.5	0.5
7	ORCID	0.0	0.0
TOTAL		2.0	1.5

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

2.3. Project management during the period

The main tasks of WP1 (Project Coordination) and their achievements are summarized below.

Task 1.1: Administrative and Financial co-ordination, M1 - M24 (BL)

Although financial reporting period is not due until the end of the first year, the Project Coordinator outlined the obligation and responsibilities of each partner to keep records of expenditure and resources used in the project at the 1st project meeting in October 2012. The pre-financing was received from the Commission and distributed according to the Consortium Agreement.

Task 1.2: Project Quality Plan, M1 (BL)

The Project Quality Plan includes provisions such as internal reporting procedures, configuration management arrangements, standard project document templates and use of management tools. The Plan was produced and submitted as deliverable D1.1 Project Quality Plan and Risk Register.

Task 1.3: Define impact metrics, M1 (BL)

A project plan with deadlines for project milestones and deliverables and dependencies was developed in M1 and fed into the Project Quality Plan. This project plan is available on the project intranet and constantly monitored.

Task 1.4: Risk register, M1 – M24 (BL, DATACITE, CERN, ANDS)

The risk register was put in place and available on the project intranet and constantly monitored. All partners have access to it and have the possibility to edit and/or included. An initial risk register was submitted as part of D1.1.

Task 1.5: Consortium management, M1 – M24 (BL, DataCite)

Records of consortium management are kept on the project extranet accessible to partners. There is regular communication between the Consortium Members via email and fortnightly telephone calls. Face to face meetings between the European partners are held at a minimum frequency of 6 months. Effort is recorded for each partner and work package on a monthly basis through the project extranet. Templates have been created for deliverables and reports to ensure compliance with the project quality plan.

Deviations from DoW/ impact on other tasks

No particular deviations have been observed for this WP.

	D1.2 First interim scientific report		
	WP1: PROJECT MANAGEMENT		Dissemination level: PU
	Authors: Sergio Ruiz (DataCite), John Kaye (British Library)		Version: 2_0 Final

Use of resources (including actual vs. planned)

Table of person-months per partner compared with the DoW.

No.	Short name	PM expected to be spent during the first 6 months of the project	PM spent in WP1
1	ANDS	0.0	0.0
2	arXiv	0.0	0.0
3	BL	2.5	1.8
4	CERN	0.3	0.1
5	DataCite	0.8	1.7
6	DRYAD	0.0	0.0
7	ORCID	0.3	0.0
TOTAL		3.8	3.5

2.4. Deliverables and milestones tables

2.4.1. Deliverables

TABLE 1. DELIVERABLES										
Del. no.	Deliverable name	Version	WP no.	Lead beneficiary	Nature	Dissemination level ¹	Delivery date from Annex I (proj month)	Actual / Forecast delivery date Dd/mm/yyyy	Status No submitted/ Submitted	Comments
D1.1	Project Quality Plan & Project Risk Register	1.0	1	BL	Report	Public	09/12	01/13	Submitted	
D2.1	Kick-off preparation, Communication plan and Website	1.0	2	ORCID	Report	Public	09/12	12/12	Submitted	
D2.2	Kick-off report	1.0	2	ORCID	Report	Public	09/12	01/13	Submitted	

PU = Public

PP = Restricted to other programme participants (including the Commission Services).

RE = Restricted to a group specified by the consortium (including the Commission Services).

CO = Confidential, only for members of the consortium (including the Commission Services).

Make sure that you are using the correct following label when your project has classified deliverables.

EU restricted = Classified with the mention of the classification level restricted "EU Restricted"

EU confidential = Classified with the mention of the classification level confidential " EU Confidential "

EU secret = Classified with the mention of the classification level secret "EU Secret "

ODIN is co-funded by the EC under the e-Infrastructures Activity of the FP7 Capacities Specific Programme

© 2013 ODIN Consortium. Some rights reserved. This work is licensed to the public under the Creative Commons Attribution 3.0 License. <http://creativecommons.org/licenses/by/3.0/>



2.4.2. Milestones

TABLE 2. MILESTONES							
Milestone no.	Milestone name	Work package no	Lead beneficiary	Delivery date from Annex I dd/mm/yyyy	Achieved Yes/No	Actual / Forecast achievement date dd/mm/yyyy	Comments
MS1	Agreements concluded	WP1	BL	30/09/2012	No	Forecast 03/13	All partners acceded to the consortium agreement with the exception of ANDS. This delay is due to the Southern Hemisphere Summer Holiday season and signature is now in progress.
MS2	Kick-off complete	WP2	ORCID	30/11/2012	Yes	10/12	D2.2 Kick-off report submitted

ODIN is co-funded by the EC under the e-Infrastructures Activity of the FP7 Capacities Specific Programme

© 2013 ODIN Consortium. Some rights reserved. This work is licensed to the public under the Creative Commons Attribution 3.0 License. <http://creativecommons.org/licenses/by/3.0/>

