

An Interoperability Infrastructure for Digital Identifiers in e-science

Barbara Bazzanella¹ and Paolo Bouquet²

¹ University of Trento, Italy
barbara.bazzanella@unitn.it

² OKKAM srl, Trento, Italy
bouquet@okkam.it

Abstract. The rapid increase of scientific digital assets in the last years has made clear that digital identifiers are crucial for effectively publishing, accessing and managing digital information in e-science contexts. From persistent keys for access to digital objects in network environments, the concept of persistent identifiers has been more recently extended to identify also physical objects like people, institutions and any type of relevant entity in the e-Science domain, opening the way to the creation of an integrated information space where a network of resources can be resolved, linked, navigated and analyzed, as the Linked Open Data approach envisions for the Web. However, the creation and full exploitation of this valuable network of connections is currently hindered by the fragmentation and lack of coordination of the digital identifier ecosystem. The aim of this paper is to propose an open, distributed and scalable infrastructure for interoperating existing Persistent Identifiers and other digital identifier systems (like Cool URIs) in e-science, overcoming geographical, disciplinary and organizational boundaries. The Digital Identifier interoperability infrastructure is presented as a cross-cutting solution of core services enabling interoperability at three different levels: identifier, co-reference and semantic.

1 Introduction

In the last 20 years, Internet and the application of computer technology to the scientific production and dissemination have contributed to change the nature of the scientific research, promoting the transition from science to e-science and posing new challenges in terms of preservation, management and sharing of digital scientific content. Digital Identifiers are at the core of this transformation and they have been largely invoked as fundamental elements for realizing the potential of e-science by providing global keys for information access, reuse and exchange and creating a complex network of connections among the relevant entities in the research data landscape (e.g. linking publications to authors and datasets, authors to institutions and projects, projects to research products and fundings). The creation and full exploitation of this valuable network of connections is currently hindered by the fragmentation and lack of coordination of the digital identifier ecosystem.

On the one hand, the community of librarians, publishers and public administrations have developed and implemented models and systems for assigning Persistent Identifiers (PIs) to digital and non-digital resources, like electronic documents and authors. However, the lack of interoperability between PIs and related vocabularies, together with the creation of multiple identifiers from different systems for the same objects³, represents an obstacle for the creation of services that manage and integrate data over multiple systems and is the first major hurdle for the development of a globally connected e-science landscape.

In parallel, the WWW has increasingly evolved into a platform for publishing data (including user-generated data and Semantic Web data) becoming a virtually infinite publication space where each resource is uniquely identified by a web name (URI) and made retrievable through standard protocols. In this context, the Linked Data community, with its concept of Cool URI, is emerging as a potential different approach to manage identifiers for digital objects, authors and other e-science entities (using the Web as a platform). Despite some criticisms concerning persistence, authority and trust issues, an increasing awareness⁴ is emerging in part of the PI community that the Linked Data practices and tools may offer a way for extending the value of data (in particular, through cross-linking) and cover use cases which traditional solutions were not designed to address. However, currently an integrating solution for harmonizing and coordinating the two approaches is far to be realized and this lack of integration represents a second crucial interoperability issue for the realization of a navigable digital identity space within the e-science landscape.

The aim of this paper is to propose an open, distributed and scalable infrastructure which both enables the interoperability among current PI systems, minimizing the drawback of the co-existence of multiple PI systems for the same objects and at the same time creates a bridge between PIs and other digital identifiers, like Cool URIs to address the challenge of managing persistent identifiers on a vast scale in open distributed environments and across systems boundaries. The Digital Identifier interoperability infrastructure is presented as a thin layer of core services enabling interoperability at three different levels: identifier, co-reference and semantic. These core services are conceived as essential building blocks for developing value-added e-infrastructure services such as services for 1) data and information discovery and navigation, 2) data sharing and linkage, 3) reputation assessment and citation.

³ An author for example may have multiple identifiers from different systems, like discovery services (e.g. Google scholar profile, AuthorClaim, ORCID), libraries (VIAF, ISNI), disciplinary systems (arXiv, PubMed Author ID), publishers (Scopus ID, Researcher ID) or social networks (LinkedIn, Research Gate), and this redundancy makes difficult to link the author to his/her research activities and scholarly work, having strong consequences in terms of discovery, reputation and impact tracking.

⁴ Several initiatives like the Persistent Object Identifiers seminar at The Hague in June 2011 and the Links That Last workshop in Cambridge in July 2012 have highlighted the need of developing a co-ordinated solution to identifier issues across the PI and the Linked Data community, as stated for example in the Den Haag Manifesto available at <http://www.knowledge-exchange.info/Default.aspx?ID=462>

1.1 Research and Initiatives on Identifier Interoperability

In the last few years a number of initiatives and projects have started to address the problem of (persistent) identifier interoperability as one of the big issues for the realization of global information infrastructures for science and cultural heritage. One of the first studies on this topic, named DIGOIDUNA [5], was conducted on behalf of the EC in 2011. The study has investigated the fundamental role of digital identifiers as enablers of value in e-science infrastructures and has performed a detailed analysis of strengths, weaknesses, opportunities and threats of the current digital identifier landscape in order to identify the main challenges and propose a set of recommendations which policy makers and relevant stakeholders should address to develop an open and sustainable persistent identifier infrastructure supporting information access and preservation. One of the main conclusions of the study is that to transform digital identifiers from simple means to manage data to keys for delivering value to the stakeholders within the research production, it is necessary to foster the development of an interoperable, cross-domain infrastructure for persistent identifiers supporting data access and sharing across national, organizational, disciplinary and technological boundaries. The implementation of this infrastructure poses not only technical challenges but raises a multidimensional spectrum of organizational, social and economical issues which should be addressed to ensure a coordinated ecosystem. A recent effort in investigating this multifaceted set of issues with the purpose of harmonizing the current identifiers solutions has been made within two working groups of the Research Data Alliance (RDA), an international forum focused on solutions for enabling open sharing of data. The RDA PID Interest Group⁵ aims to bring together relevant stakeholders and practitioners to define emerging PID use cases in the domain of data and coordinate the use of persistent identifiers for supporting referencing and citation of research products and authors, and manage the lifecycle of research data production. The PID Information Types Working Group⁶, is focused on the definition of a common framework of information types to be associated with PIs for a proper data management and access.

In line with the DIGOIDUNA idea that identifier interoperability should be built on the social and organizational complexities of the current solutions, some initiatives have started to define cooperation agreements and complementary architectures to ensure interoperability between independent systems or organizations. ORCID and ISNI for example have agreed to render ORCID compatible with the ISNI ISO standard and assigning a block of numbers for identifying ORCID entities which can not be reassigned by ISNI to different people⁷. In addition, ORCID has entered into an agreement with Ringgold to use it as Registration Agency for ISNI Institutional Identifiers to support the persistent identification of researcher institutional affiliations in the ORCID reg-

⁵ <https://rd-alliance.org/internal-groups/pid-interest-group.html>

⁶ <https://rd-alliance.org/groups/pid-information-types-wg.html>

⁷ <http://orcid.org/blog/2013/04/22/orcid-and-isni-issue-joint-statement-interoperation-april-2013>

istry. The integration between Researcher ID and ORCID is another example of a bi-directional integrating initiative aimed at making information on the two systems interoperable and complementing. Similarly, the ODIN project⁸ aims to define a roadmap for the integration and scalability of the DataCite and ORCID identifiers solutions to create a layer of interoperability between persistent identifiers for researchers, research works and their outputs (publications and data). The proposed solution is based on a conceptual model of interoperability [3] for linking research data and their contributors (embedding the corresponding PIs into metadata) through the coordination and alignment of the information flow across data centers, DataCite, and ORCID.

These solutions have made a concrete step forward in the coordination process within the identifier landscape but they have the limit to define interoperability in a point-to-point way. A more broad and integrated approach can be found in the APARSEN project. Within APARSEN, the research on persistent identifiers has focused mainly on the definition of an interoperability framework for persistent identifier systems [1] which defines some key assumptions and requirements to identify the trustable candidate systems which can take part to the framework, an ontology which specifies the structure of data and the core set of relationships linking the identified entities within the framework and finally a small set of services which can be implemented on top of the framework. A demonstrator⁹ has also been developed to provide evidence of the potential applicability of the model and the value of related basic services [2].

Finally, other initiatives have been started within specific communities. In the library domain, the BIBFRAME initiative¹⁰ has defined a lightweight framework (metamodel) for bibliographic description based on linked data principles to improve the integration, discoverability and reuse of library resources and their descriptions in a networked distributed environment. At the core of the model, there is the concept of BIBFRAME authority which is a resource representing a person, organization, place, topic, temporal expression and other entities associated with a BIBFRAME Work, Instance, or Annotation. BIBFRAME authorities are used non only to identify (via URIs) the above mentioned entities within the description, but also to link to external resources (for example traditional authorities) referring to the same entities by including their corresponding IDs. In this way, the mechanism of BIBFRAME authorities should provide a common lightweight interoperability layer over different Web based authority resources connecting a BIBFRAME resource to one or more authorities for related entities, such as a person, organization, or place, identified by other identifiers systems like a ID.LOC.GOV, ISNI, VIAF and others. All these initiatives have the merit to have increased the awareness and consensus among the relevant stakeholders and communities about the crucial role of a coordinated ecosystem of persistent identifiers at the heart of a global infrastructure for e-science. However, a solid technological solution for interoperating

⁸ <http://odin-project.eu/>

⁹ available at <http://93.63.166.138/demonstrator/demo7/>

¹⁰ <http://www.loc.gov/bibframe/>

identifiers for digital objects, contributors, authors and other relevant entities is still a lacking aspect in the effort of developing a sustainable infrastructure providing an invisible layer of interoperability on which cross-cutting advanced services for science and education can be implemented. This is pointed out in the recent EU Framework Programme for Research and Innovation (Horizon 2020)¹¹ which includes the implementation of a Digital-identifier e-infrastructure for digital objects, contributors and authors, among the key actions for implementing an open, interoperable e-infrastructure for scientific data. Based on the valuable results of the above initiatives, but also exploiting the experience on global identifiers for the Semantic Web gained in the course of the OKKAM FP7 project¹², this paper proposes a technical solution to address this challenge by implementing an interoperability cross-cutting system for Persistent Identifiers and other identifiers used in e-science.

2 Implementing an interoperability infrastructure for digital identifiers

The infrastructure that we propose is an open, distributed, decentralized and scalable service for managing unique identification of digital and non-digital resources over digital networks and implementing interoperability services for existing identifier systems, including PIs and Cool URIs. The implementation of the infrastructure has been driven by the following requirements:

- **IDENTITY MANAGEMENT:** the infrastructure is aimed at managing the unique identification of digital, physical or abstract entities of any type (e.g. person, organization, event, artifact, location), where there is a need to distinguish them from other objects in digital content and interoperate all the external identifiers used to identify the same entities. It is important to keep separate the ID management from the management of the information identified by the ID. While the interoperability infrastructure should perform ID management, in our approach content management should be handled outside the infrastructure by data managers and content providers. However, since different identifiers use different identifier schemes to identify the same objects, the ID management should ensure not only interoperability at the level of identifier (e.g. connecting equivalent identifiers from different sources) but also interoperability at the level of the identifier metadata which includes the defining attributes that are considered sufficient to establish the unique reference between communicating parties. ID management should also be kept separated from the implementation of value-added community services. The interoperability infrastructure is conceived as a thin layer of core ID management services on top of which advanced services can be built to address specific needs and requirements of specific users. These

¹¹ http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/main/h2020-wp1415-infrastructures_en.pdf

¹² <http://project.okkam.org/>

value-added services are not part of the infrastructure but are enabled by the infrastructure and can be developed by third parties and driven by targeted business models.

- PERSISTENT COOL IDs: the identifier solution implemented by the system should be compatible both with persistence and authority requirements of the PI users and with openness and decentralization principles of the Web community.
- OPENNESS: the infrastructure should not present entrance barriers for existing identifier providers and users which are encouraged to build their applications and services on the infrastructure.
- DISTRIBUTION and DECENTRALIZATION: the infrastructure operations and data can be distributed and replicated over any number of different machines both geographically and logically in a seamless way. Moreover, the infrastructure must have no single point of failure and it must be able to survive any damage in any node of its architecture. This redundancy is crucial for ensuring the PERSISTENCE and RELIABILITY of the service. The distributed and decentralized architecture contributes also to the SUSTAINABILITY of the system whose costs can be shared and distributed among multiple organizations hosting the nodes of the infrastructure.
- SCALABILITY: the infrastructure should be able to cope with billions of digital identifiers from any number of providers and serve hundreds of inter-operability requests per second.
- TRUST: The infrastructure is conceived as a trusted infrastructure where an operating organization is committed to the long term (i.e. Trustee) under the surveillance of an international board of protectors representing its main stakeholders.

The proposed infrastructure is the result of the evolution of the Entity Name System (ENS),¹³ a scalable Web service, developed in the context of the EU-funded OKKAM project¹⁴ for assigning and managing unique identifiers for Web entities and foster their global reuse. The first prototype of this system has emerged as a solution for the entity identification problem in the Semantic Web [6], that is the problem of integrating information about entities which are assigned multiple identifiers (i.e. URIs) in different systems or by different users [7]. Recently, the scope of the system has been extended to address the problem of the proliferation of identifiers (i.e. PIs) in other distributed management information contexts like libraries, archives, publisher systems and e-science infrastructures, evolving into a bridge infrastructure for Web identifiers and PIs.

The first aim of the ENS is to provide a service to assign globally unique identifiers to entities named in information sources and reuse these identifiers across systems boundaries regardless of the place or domain where they have been first

¹³ <http://api.okkam.org/>

¹⁴ <http://project.okkam.org/>

assigned. To this purpose, the ENS has a scalable repository¹⁵ for storing entity identifiers along with a short set of descriptive metadata, i.e. an entity profile, which is used to disambiguate each entity from the others. When a human user or an application inquire the system for an identifier (for example by keywords through the Web search interface¹⁶), the information of the entity profiles is used to establish (through advanced entity matching algorithms) if an identifier has been already assigned and stored for that entity. Otherwise, a new identifier is minted and returned by the system. The systematic reuse of the identifiers created and maintained in the ENS is aimed at reducing the multiplication of identifiers for entities and enabling a straightforward solution for entity-centric integration of information spread across different systems and platforms. However, since the convergence of different users on the use of a unique identification standard is desirable but very hard to be realized at a global level, the coordination and interoperability of different identifier systems has emerged as the second main function implemented in the ENS. Therefore, the ENS implements a service to store (within the entity profiles) and manage equivalent identifiers (i.e. alternative IDs) assigned to the same entities by different systems. This service, called Mapping service, together with the other services described in Table 1 support the core operation of the ENS. Local private nodes (replicas) of the infrastructure can be implemented to include the same core services with the aim of managing entities or sensitive information which need to be maintained physically separated from the rest of the public infrastructure. Today, more than 9 millions of entities are managed by the system and the number is growing every day. In the next section we will discuss how on top of this core infrastructure, the ENS has been evolved into an interoperability infrastructure bridging digital identifiers in use in e-science and other digital contexts. We name this infrastructure "Entity Identifier Bridge".

2.1 From the Entity Name System to The Entity Identifier Bridge

The Entity Identifier Bridge (EIB) provides interoperability services at three different levels:

1. The identifier level: offering an identification solution which is compatible both with the persistence and authority requirements of the PI community and with the openness and decentralization principles of the (Semantic) Web community.
2. The co-reference level: supporting cross-linkage of different PIs and other digital identifiers for the same object (identifier mapping).
3. The semantic level: storing mappings between the different metadata / vocabularies which are associated with each identifier in different systems.

¹⁵ The storage and access components of the infrastructure are based on Apache Hadoop, a scalable software framework for distributed storage and processing of big data.

¹⁶ available at <http://api.okkam.org/search/>

Service	Description	Function
STORAGE	A large-scale ENTITY REPOSITORY which stores identifiers, entity profiles and other information about different types of entity	To ensure the uniqueness and persistence of the binding between an entity identifier and the entity
MATCHING	A set of ALGORITHMS to decide whether an ID already exists, given the information a user or application presents to the system	To make identifiers searchable and easily retrievable by humans and applications
LIFECYCLE MANAGEMENT	MAN- to support few BASIC OPERATIONS on entity profiles and ENS IDs (create, edit, merge, split..) and deal with evolving entity	To ensure adequate maintainability and the secure long-term evolution of the system
ALTERNATIVE MANAGEMENT	ID to store known MAPPINGS between the ENS ID and any other available ID for the same entity	To enable identifier interoperability

Table 1: ENS core services

The identifier Level As discussed in [4], one of the first actions that contributed to evolution of the ENS into the EIB is the change of the ENS identifier syntax to offer a reconciling identification solution which can be used both by PI communities to implement trusted authority-based services and by the Semantic Web and Linked Data users to publish and distribute structured content on the Web. The simple change consists of the separation of the ENS-ID (e.g. `ens:eid-a5afe3ae-071f-4ec3-b904-3481aa5a6a05`) from the resolver (`http://www.okkam.org`). This introduces a level of indirection between the identifier and its referent, decoupling persistent identification from resolution and retrieval. When the ENS-ID is combined with the ENS default resolver, its resolution returns a small set of metadata (included in the ENS entity profile) related to the identified entity. The real potential of separating the token ID from the resolver rests on the possibility of associating the same ID to multiple resolvers, enabling a mechanism of multiple resolution and distribution of authority in the management of data. Different actors can create or reuse Persistent ENS-IDs for entities of interest using the ENS and through their local resolvers or domain names enable precise (and possibly long-term) access to information they want to provide. In this way, ID management and default resolution are addressed by the ENS, whereas information management, including persistence of content and reliable resolution, is managed by content providers. The ENS-IDs can be used to create Cool URIs (through the web resolution protocol and the DNS) allowing Linked Data users to create URIs resolvable to any information source they like. At the same time, persistent identifiers users can reuse the same IDs to identify information objects and resources managed by trusted institutions which ensure their persistent access and association to a physical location. This improvement of the ENS may offer a co-ordinated solution to identifier

issues across the PI and the Linked Data communities, enabling data creators and curators to combine the technical strengths and opportunities of the (Semantic) Web vision with the organizational, economical and social requirements legitimately raised by PI stakeholders with the result of opening new forms of interactions between open structured data published on the Web and content stored and preserved by more traditional cultural heritage institutions.

The co-reference Level At this level the EIB implements a small set of core interoperability services on top of the Mapping Service of the ENS to enable access and reuse of mappings between alternative identifiers for the same entity. Currently three services have been implemented. The first two services support the retrieval of alternative IDs. The third service implements the multiple resolution of entity identifiers.

Alternative ID Retrieval Service (by ID): The first service allows to retrieve an alternative ID of a specific type (which can be a PI or Cool URI) or all the available alternative IDs for a given entity ID (which can be a PI or Cool URI). The user for example can enter the ISNI of a person to get the ORCID ID of the same person (as shown in Figure 1) or retrieve all the alternative identifiers stored in the system for her.

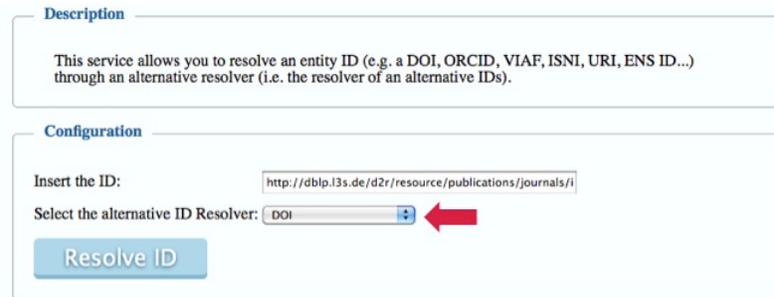
The image shows a web interface for the Alternative ID Retrieval Service (by ID). It is divided into two main sections: Configuration and Results. In the Configuration section, there is a text input field labeled "Insert the ID:" containing the URL "http://isni.org/isni/0000000047782063". Below it is a dropdown menu labeled "Select the alternative ID Type:" with "orcid" selected. A red arrow points to the dropdown menu. A blue "Get ID" button is positioned below the dropdown. In the Results section, the text "orcid:" is followed by a red-bordered box containing the URL "http://orcid.org/0000-0003-2811-0896".

Fig. 1: Alternative ID Retrieval Service (by ID)

Alternative ID Retrieval Service (by Query): The second service supports the same function as the previous one with the only difference that the identified entity is searched by a keyword query. Going back to the previous example, the user can enter a keyword query (e.g. the name of the person) to get an alternative ID of a specific type (e.g. the ISNI) or all the available alternative IDs for that entity.

Alternative Resolution Service: The last service allows to enter an ID and resolve it through the resolver of an alternative ID. For example, the user can enter the Cool URI of a resource (as shown in Figure 2) and access to the

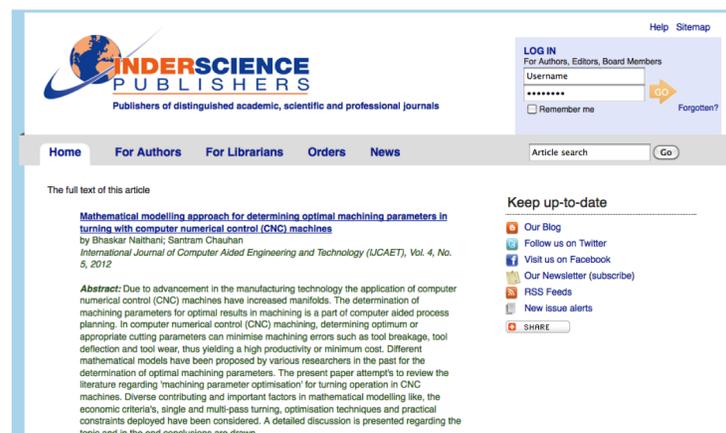
resource (or a description of it) through the resolver of a selected alternative ID. In the example in Figure 2 the selected alternative ID is the DOI of the resource which resolves to the InderScience Publishers page for the identified article.



The screenshot shows a web form with two main sections: "Description" and "Configuration".

- Description:** A text box containing the text: "This service allows you to resolve an entity ID (e.g. a DOI, ORCID, VIAF, ISNI, URI, ENS ID...) through an alternative resolver (i.e. the resolver of an alternative IDs)." This section is enclosed in a light blue rounded rectangle.
- Configuration:** A section containing:
 - An input field labeled "Insert the ID:" with the text "http://dblp.l3s.de/d2r/resource/publications/journals/i" entered.
 - A dropdown menu labeled "Select the alternative ID Resolver:" with "DOI" selected. A red arrow points to this dropdown.
 - A blue button labeled "Resolve ID".

(a) Input



The screenshot shows the homepage of InderScience Publishers. The header includes the logo and tagline "Publishers of distinguished academic, scientific and professional journals". A navigation menu contains links for Home, For Authors, For Librarians, Orders, and News. On the right, there is a "LOG IN" section for authors, editors, and board members, with fields for Username and Password, a "GO" button, and a "Remember me" checkbox. Below the navigation is an "Article search" field with a "Go" button. The main content area displays the full text of an article titled "Mathematical modelling approach for determining optimal machining parameters in turning with computer numerical control (CNC) machines" by Bhaskar Nalithani and Santram Chauhan. The abstract discusses the application of computer numerical control (CNC) machines and the determination of optimal machining parameters. On the right side, there is a "Keep up-to-date" section with links to the blog, Twitter, Facebook, newsletter, RSS feeds, and new issue alerts, along with a "SHARE" button.

(b) Output

Fig. 2: Alternative Resolution Service

The Semantic Level At the third level, identifiers are linked across meta-data/vocabularies. Different persistent IDs are usually associated to different metadata models. If a mapping among them is available, the information represented according to a data model in one source can be translated into a different model and re-used to integrate or update the information of another source adopting that model. Therefore, in order to support semantic interoperability across services and communities, the ENS should provide an extensive mapping of vocabularies and schemes adopted in different PID domains. In addition, since these mappings can be also defined contextually and different users can have their

own mappings for the same vocabularies, a solution for managing these mappings and making them reusable, should be the result of a large-scale social process where users can create their mappings but also reuse the mappings defined by other users. Following this idea, a Web service, called OKKAM Synapsis¹⁷, has been developed to store and maintain user-defined mappings between terms in controlled vocabularies, models and ontologies. Through the application, a user (human user or API user) can search the available mappings for the classes and properties of the models present in the system, add new mappings, rate (like it) and edit existing mappings (e.g. tagging, adding a description). In addition, the user can upload a new model (i.e. the target ontology) and define his/her own mappings for the model. The mappings are defined by specifying the Resource URI, that is the URI of the resource mapped towards the element of the target ontology and the type of relation between them¹⁸. The service allows to filter the stored mappings by author, tags, status and type of relation. While in the APARSEN Interoperability Framework semantic interoperability is addressed by proposing a common ontology which should be used by content providers to expose their data in a common way, the ENS approach focuses on the alignment of different vocabularies through ontology mapping. This has the advantage that users can maintain their own vocabularies and ontologies, without the need of restructuring their content according to a new model. The mapping of vocabularies allows to support the building of crosswalks between them and can be extended to include an indefinite number of vocabularies.

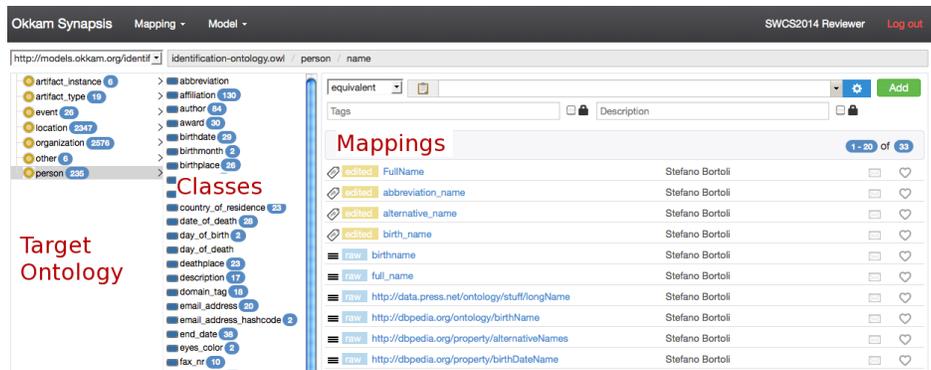


Fig. 3: Synapsis Interface

¹⁷ <http://api.okkam.org/synapsis/>

¹⁸ the user can select among a number of relations types including OWL meta-relations such as owl:EquivalentProperty, owl:EquivalentClass, owl:SubClass, owl:SubProperty and SKOS meta-relations skos:exact, skos:close, skos:broader, skos:narrower.

3 Conclusion

In this paper, we propose an open scalable interoperability infrastructure (EIB) for persistent identifiers and other digital identifiers in e-science. The proposed infrastructure is based on the OKKAM Entity Name System and is presented as a thin layer of core interoperability services which enable interoperability 1) at the level of identifier, providing a reconciling identification solution for PIs and Cool URIs, 2) at the level of digital identifier co-reference, implementing services for alternative id retrieval and resolution and 3) at the semantic level, providing a service for managing, sharing and reusing user-defined mappings across identifier data models. On top of the core services of the infrastructure value-added services can be built which can take advantage of the interoperability layer to resolve, navigate and link entities across digital curation, discovery, collaboration, authentication and other e-Infrastructures services and extract value from distributed data. For example, author profile management services can be built to automatically import, integrate and synchronize information from multiple sources like authoritative sources (e.g. libraries and publishers), auto-claiming services or Web content and trace the provenance of the information. Empowered information discovery services can provide multiple access to digital resources stored in different systems via alternative IDs or provide integrated views of these resources through the mash-up of metadata from different sources using different schemes. Discovery and navigation tools can be built which allow to go through the chain of links connecting the identified entities in the e-science landscape (e.g. authors-publications-datasets-projects-fundings), discover hidden relationships and generate new valuable knowledge (e.g. who contributed in what ways to the research outputs with the support of which fundings and institutions.).

References

1. APARSEN D22.1: Persistent identifiers interoperability framework. http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/04/APARSEN-REP-D22\textunderscore_1-01-1\textunderscore_9.pdf, 2012.
2. APARSEN D22.3: Demonstrator set up and definition of added value services. http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2014/02/APARSEN-REP-D22_3-01-1_0.pdf, 2013.
3. D4.1: Conceptual model of interoperability. http://files.figshare.com/1239137/D4.1_Conceptual_Model_of_Interoperability.pdf, 2013.
4. B. Bazzanella, S. Bortoli, and P. Bouquet. Can persistent identifiers be cool? *IJDC*, 8(1):14–28, 2013.
5. P. Bouquet, B. Bazzanella, M. Dow, and R. Riestra. DIGOIDUNA FINAL REPORT: Digital Object Identifiers and Unique Author Identifiers to enable services for data quality assessment, provenance and access. <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/digoiduna.pdf>, 2011.
6. P. Bouquet, H. Stoermer, and B. Bazzanella. An entity name system (ens) for the semantic web. In *ESWC*, pages 258–272, 2008.
7. P. Bouquet, H. Stoermer, C. Niederée, A. M. Na, C. Niederee, and A. Mana. Entity Name System: The Backbone of an Open and Scalable Web of Data. In *ICSC*, number CSS-ICSC 2008-4-28-25, pages 554–561. IEEE Computer Society, 2008.