

# Art and design as linked data: the LODZ project (Linked Open Data Zurich)

**Nicolas Prongué, nicolas.prongue@hesge.ch**

University of Applied Sciences HEG Geneva, Switzerland

**Fabio Ricci, fabio.ricci@semweb.ch**

Semweb LLC, Küsnacht, Switzerland

**René Schneider, rene.schneider@hesge.ch**

University of Applied Sciences HEG Geneva, Switzerland

**René Schurte, rene.schurte@zb.uzh.ch**

Zentralbibliothek Zürich, Switzerland

Libellarium, IX, 2 (2016): 187 – 202.

UDK: 029.49:004=111; 7.025: 004.62=111

DOI: <http://dx.doi.org/10.15291/libellarium.v9i2.256>

Research paper

## Abstract

The project LODZ (Linked Open Data Zurich) adopts an experimental approach to merge data and develop a semantic web infrastructure to enable its discovery. For this purpose, three institutions in the field of art and design provided their metadata. The project cycle followed six steps: team building, gathering and cleaning of the original data, modelling, transforming, interlinking and exploration of the Linked Data set. The resulting pilot application offers innovative and attractive features based on the capability of the Linked Data, with the aim to provide a better user experience.

The major challenge of this project was the creation of links between the internal datasets, and with external sources. An important lesson learnt is therefore to focus more on the interoperability of data at the time of cataloguing in the original databases, for example by integrating external identifiers rather than just terms in the form of strings.

**KEYWORDS:** application, cultural data, linked data, semantic web

## Introduction

### Background

For centuries the cultural institutions of our societies gathered objects and created high quality metadata to describe them. These objects were mainly collected in and for the public sector, i.e. in libraries, archives and museums, by creating their own informational universe. Specific formats, specific cataloguing

rules, specific protocols and specific data models were developed in each area, making it easy to exchange data within this area, but difficult to use and understand by external systems. The most glaring example of this dichotomy is that of libraries, with their complex MARC format, the Z39.50 protocol and the AACR2 rules (Bermès 2013).

Besides that, the modern digital society has gone and is still going through several revolutionary and evolutionary transformations that have not stopped outside the doors of the institutions mentioned above. Data were first recorded on paper, then digitized and later structured and stored in databases. Finally, the disruptive fracture caused by the emergence of the World Wide Web has led to a new paradigm: the automatic exchange of information, beyond the borders of a closed system. Data were consequently serialized in XML and nowadays, for the purpose of interlinking, converted according to the standards of the semantic web or better Linked Open Data, defined as data published using the Resource Description Framework (RDF) (Coyle 2016).

Linked Data also became a W3C standard, i.e. that each statement is made in a triple structure composed of a subject, a predicate and an object. The resources are identified by HTTP URI and linked with external data. They are often coupled with an open licence, turning them into Linked Open Data (Berners-Lee 2010, Europeana 2015). Since its standardisation in 2004, the semantic web has expanded to form a huge data cloud of more than one thousand datasets, interlinked around an important node, the DBpedia dataset (Schmachtenberg, Bizer and Paulheim 2014).

The web user requirements are increasingly high: online data services have to offer more serendipity and have to be visible outside of their closed institutional environment. Interlinking the data through the RDF standards is a potential answer to these needs, nevertheless, it also represents a significant technical challenge; indeed, it implies various tasks that still nowadays are considered complex, like the full automation of data processing, the attribution of persistent identifiers or the quality assessment (Bensmann 2016). These challenges are exacerbated by problems concerning legacy data and legacy systems, especially in public institutions.

Due to the fact that cultural data is mainly kept in public institutions as mentioned in the initial sentences of this paper, the Linked Data movement opens a road to libraries, archives, and museums toward the convergence and reshaping of their cultural data.

### **Related works**

Libraries are the most advanced public institutions in this area. Some of them have made their data available in RDF via web services (API) or dump, but only a few of them have already created a Linked Data application for the end user. Among them is the French National Library and its project *data.bnf.fr* with the aim of merging various internal datasets, interlinking them with each other

and providing a search interface based on the RDF framework. The resulting product has become one of the most advanced semantic web applications in libraries, presenting a clear evolution compared to traditional library catalogues: FRBR-structured data, new easy-to-use features and design patterns based on enriched data, links to external datasets as well as data openness with download possibilities (BNF 2016).

In the area of digital libraries, some initiatives have resulted in successful Linked Data applications. Europeana developed an RDF-based model, which enables a unified representation of digital objects gathered from heterogeneous data providers. The Europeana Data Model distinguishes the cultural heritage object from its online representations and its metadata (Charles 2016). Based on this framework, the website *Deutsche Digitale Bibliothek* of the German National Library is a working application which proposes interesting Linked Data features, like pages for persons with links to other external datasets (Wikidata, ISNI, Library of Congress, etc.).

Other projects have been more focussed on cross-domain contents from GLAM institutions (galleries, libraries, archives and museums). Kulttuurisampo merges and links all kind of GLAM data (objects, texts, pictures, places, etc.) from all over Finland, and provides it through a single web interface (Mäkelä, Hyvönen and Ruotsalo 2012). Significant efforts have been made in the creation of new functionalities to explore the RDF graph, breaking away from the paradigm of the traditional search engine with a search bar and a results page. In this context, it must be mentioned that an ordinary web user may not be able to benefit from all these advanced functionalities without assistance.

In France, the creation of the new website for the Centre Pompidou followed the same approach at a smaller scale, within a single institution. Combining data from the library, the archive, and the museum as well as other in-house data, the information has been remodelled in RDF around core concepts that appear in all data sources, such as events and people (Dalbin et al. 2011). The website presents original interlinked contents, such as works of art, online educational dossiers or even products of the museum shop, in a very user-friendly interface.

These examples show the search of a different, in the ideal case better user experience, by providing well-interconnected information and mash-up services on the web. However, such mature projects need the firm and consistent investment of a large institution to ensure stability, durability, and sustainability.

### **The LODZ project**

The project LODZ (Linked Open Data Zurich), launched in Switzerland in the spring of 2015, adopts an experimental approach to merge data around art and design. The aim of the project was firstly to get a thorough overview of the convergence of cultural metadata and their conversion to Linked Data, in order to determine a common workflow for data transformation. Secondly, this knowledge was applied to a concrete situation, mostly to gain practical

experience, by the development of a pilot application which merges and interlinks heterogeneous datasets in RDF and proposes innovative search features to explore the data. As quite a lot of mash-up services are primarily interested in specialized datasets of a particular field (Hügi and Prongué 2014), we decided to focus on the subject of art and design in the city and canton of Zurich, Switzerland.

The city of Zurich and its surrounding areas are a centre of international art trade and also a historical site in the history of art (as, e.g., the Dada movement that emerged in Zurich exactly a century ago, in 1916). The city has many important galleries and art institutions as, e.g., the Kunsthaus Zürich. Important works of art with their corresponding metadata are assembled in collections affiliated to public institutions. As is often the case in the federal structure of Switzerland, the governance of these institutions is manifold. Some are funded by the canton (as the University of Zurich), the Swiss Federation (as the Swiss Federal Institute of Technology with important collections of, e.g., graphic material), or by public or private foundations. Each institution's collection has its separate tradition of cataloguing and the different institutions have only recently started to intensify their cooperation.

From this point of view, the field of art and design in the city and canton of Zurich was a fitting example to test the benefits of the Linked Data approach. We chose several institutions willing to participate in the project, namely the Collection of Graphic Materials of the Zentralbibliothek Zürich, the Swiss Institute for Art Research with its online database of Swiss arts as well as two datasets from the Zurich University of the Arts, namely its Media Archive and the database hosted by the Museum of Design which supplied their datasets with reference to Zurich to the project.

These four collections are heterogeneous in two different aspects: they cover different fields of art and design and their databases are technically diverse. The core areas of the collections are as follows:

- The Swiss Institute for Art Research (Schweizerisches Institut für Kunstwissenschaft) is a research institute on Swiss art based in Zurich. It maintains varying online sources, among them SIKART<sup>1</sup>, an online encyclopaedia on Swiss art.
- The Collection of Graphic Materials<sup>2</sup> at Zentralbibliothek Zürich (Graphische Sammlung der Zentralbibliothek Zürich) is an art collection with focus on local and cultural history of Zurich, containing over one million items from the 15th century onwards.

---

1 Website : <http://www.sikart.ch>

2 Website: <http://www.zb.uzh.ch/spezialsammlungen/graphische-sammlung/index.html> de

- The Media Archive<sup>3</sup> of the Zurich University of the Arts (Medienarchiv der Künste der Zürcher Hochschule der Künste) is a database for the students and faculty of the University, which allows them to archive and share different kinds of material for collaborative creative work.
- The eMuseum<sup>4</sup> of the Zurich University of the Arts is the electronic archive of the University of the Arts as well as of its well-known Museum of Design (Museum für Gestaltung Zürich). It contains, i.a., a substantial collection of posters.

These four heterogeneous datasets were ideal material to test the Linked Data approach with its potential to unite search results of heterogeneous data, regardless of their technical form and their thematic background, enabling serendipitous output. The goal of our project and of the search application was to present search results on themes common to the four collections, regardless of their different focus. With such an application, it is possible to find information about a motif, an art technique or material in established Swiss art, a local historical collection, on historical posters and in works of contemporary students of art with only one request. We started on a smaller scale with representative datasets to gather experience with Linked Data in order to develop semantic features that did not exist so far on other websites.

Since the focus of the project was on the process and not on the outcome, the pilot application presented below is mainly a prototype aiming to demonstrate the benefits of Linked Data, not a definitive product.

## Methods

The development of a pilot application, which proposes innovative search features to explore the cultural data merged in RDF from different repositories, followed an approach inspired by the methods adopted in various other data conversion projects. In the domain of libraries and Linked Data, the workflows of the National Library of Spain and the LIBRIS network in Sweden, among others, are particularly relevant (HEG Genève 2015). Built on this basis, our own approach followed six – not necessarily consecutive – steps, which are detailed below: team building, gathering and cleaning of the original data, modelling the data, transforming, interlinking and exploration of the Linked Data set.

### 1. Team building

The project team comprised partners from very different sectors: the Haute école de gestion de Genève as an academic partner, the Zentralbibliothek Zürich from the public and Semweb LLC from the private sector, the latter being a company specialized in semantic web technologies. The team was then enlarged by data

3 Website : <http://medienarchiv.zhdk.ch/>

4 Website : <http://www.emuseum.ch/>

suppliers, with the Zentralbibliothek Zürich contributing its own collection of graphic materials; further data were provided by two other institutions, as mentioned in the previous sections, namely the Swiss Institute for Art Research and the Zurich University of the Arts.

As this step has no technical aspects, it might appear as a relatively easy and quick procedure. Nevertheless, the challenges are quite significant, and are located at management level. To ensure a successful development of the project, a letter of agreement had to be drawn up in collaboration with the partners and data providers. The latter committed themselves to provide the data, thus providing an indispensable prerequisite for the project. The agreement also outlines which data are concerned, over what period of time and under which conditions of use.

In this context, building a team also meant clarification of roles and responsibilities: for each data set, one person was assigned to take technical responsibility and another person to coordinate all activities. This point was decisive for the next step, namely the gathering and cleaning of the original data, which required an intensive communication between the various actors.

## **2. Gathering and cleaning of the original data**

No project dealing with Linked Data can advance without transformation of data. This fact is as simple as it is decisive and far from trivial. To make this process succeed, a strict deadline was communicated to all participants right from the start. Although some participants had no prior experience with extracting and delivering a whole database or parts of it, all data were submitted according to the deadline set after a close and intensive dialogue between the system developer and the data providers. To avoid delays, the developers accepted data in various formats. In our case the data from the four databases were delivered in four different formats, namely MARCXML, CSV, XSLX and JSON-LD, making it necessary to further process them to make them convertible.

## **3. Data modelling**

The modelling task implies a detailed and comparative analysis of the available data, and especially the identification of the common fields – possibly with different labels – in all four databases. In order to start the process of interlinking every database with each other, some specific aspects have been evaluated: the quantity of records, the type of entities described (works, people, etc.), the geographic and temporal coverage as well as the precise thematic area of the data. Controlled fields have been listed and compared; as the range of possible values is limited in these fields, a manual interlinking process would be less time-intensive and may be envisaged.

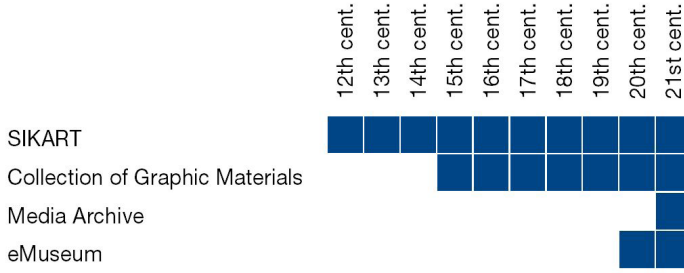


Figure 1. Temporal coverage of the four databases

This step led to the identification of significant divergences between the datasets that restrict the possibilities of interlinking (Figure 1). Indeed, the creation of links is quite difficult for databases with very different time spans covered. For that reason, considerations about potential use cases for the prototype application already started at an early stage, in order to determine on which aspect the interlinking efforts should focus. Results of this analysis showed that the interconnections should focus neither on persons and exhibitions (they are time-period-dependent), nor on places (the project already being limited to the canton of Zurich), but on general topics of interest and material types of the described resources. Further information regarding this aspect are given in section “5. Interlinking”.

Next, the modelling phase required the mapping of every relevant field in the four databases with an RDF property, to finally achieve a single homogeneous RDF model with the most detailed granularity possible. This model is illustrated in Figure 2: the two blue ovals represent the entities of interest (the subjects of our triples to be created), the arrows stand for the properties (predicates of the triples), the grey rectangles are the values in a string form (objects of the triples) and the violet ovals the values in a HTTP-URI form (being also objects of the triples). In addition to these labels, it is indicated, next to each object, in which of the four databases this value is available. Indeed, the RDF structure is flexible enough to contain a specific property only for some of the resources; for instance, the property “dct:publisher” is defined only for the works resources of the Collection of Graphic Materials. Overall the model points out where similarities between the datasets are to be found and where exactly interoperability could be achieved. Finally, a consistent internal schema of identifiers was designed for the subjects, i.e. the resources of the LODZ project.

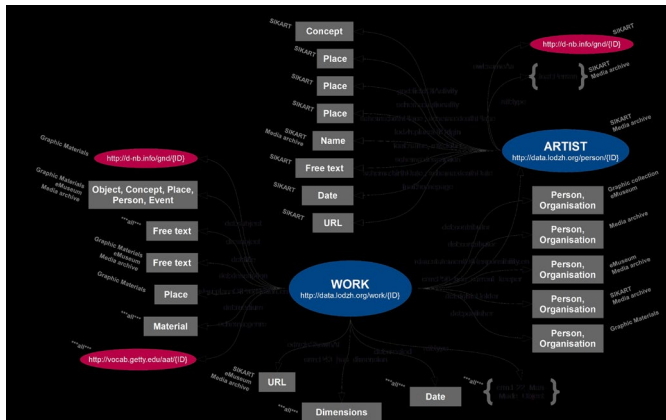


Figure 2. LODZ data model

The RDF vocabularies were chosen according to the domain of the data (works of art and documents) and their popularity on the web. Thus, the following vocabularies were used: Dublin Core (dc/dct), Resource Description and Access - unconstrained (rdau), Schema.org (schema), Europeana Data Model (edm), CIDOC Conceptual Reference Model (crm), Friend of a Friend (foaf) and GND Ontology (gnd). An additional property was defined in a small ontology created only for this project, to express the place of origin of a person: "lodz:placeOfOrigin".

#### 4. Data transformation

The fourth step, the data transformation itself, implied the setting up of a suitable infrastructure, containing an SQL/RDF mapper, tools for data analysis and conversion, and an RDF store. As some parts of the datasets were delivered in MARCXML format, they were directly converted to RDF using Metafacture, a tool designed exactly for the proper conversion of bibliographical data into RDF. The other datasets were delivered mainly as relational data and were therefore converted to RDF using a Sesame based SQL/RDF mapper developed ad-hoc by Semweb LLC.

Once the transformation infrastructure had been established, it was used to convert the data according to the model defined by the project team: for each data field in the four databases, a transformation rule was defined, either in the SQL/RDF mapper or in Metafacture, and tested on a sample of records. This transformation process is highly iterative, requiring many tests and corrections that sometimes have to be executed throughout the entire project lifecycle, depending heavily on all data requirements that are discovered during each and every step of the project.

#### 5. Interlinking

The unambiguous linkage of all datasets, the so-called *interlinking process*, may be considered to be the project's biggest challenge. The work done during

this stage was based on the same technical infrastructure as that used for data transformation.

The modelling phase made clear that these interconnections should focus on common characteristics in the four databases, especially the topics or subjects related to the resources and the material types.

To handle the topics, two thesauri were chosen: the GETTY AAT (Art & Architecture Thesaurus), mainly composed of entries in English, and the thesaurus created for the eMuseum at the Zurich University of the Arts, with all entries in German. While the GETTY AAT is already available as Linked Data, the thesaurus of the eMuseum had to be modelled and converted in RDF for the purpose of its integration into the application. This was managed using the SKOS vocabulary, a widely used semantic web ontology for the representation of knowledge organisation systems. Besides this, a module which tokenizes the search term(s) – usually a short phrase – into a list of nouns together with their stems was developed. This module allows one to use these two thesauri as a hub to cover information spread in all of the four data sources. To do so, the module subsequently goes through the list of all tokenized search terms to find on the fly the corresponding matches in the thesauri. Hence, the two thesauri adopt a hub function between the heterogeneous records enabling the user to reach all linked entities with a simple query entered as free text and to navigate among the Linked Data thesaurus.

Another interlinking operation was realized for the material types. In this case the connections were done manually by an information specialist. This operation is made possible due to a limited range of expected values in these fields. It ensures a better link quality than that an automated process would guarantee. The material types were linked with materials or contents from the GETTY AAT. This relation appears in the model under the property schema:genre. Table 1 illustrates the manual interlinking operations with one AAT identifier for the concept of *photography*; many different terms represent these same concepts in the four data sources.

Table 1. All material types that have been manually matched with one single GETTY AAT identifier

Material type as it appears in the data sources	Data source	GETTY AAT identifier
[Sachfotografie von Fleisch]	eMuseum	300046300
/ns/media_content/type/fbc6334d-42f4-40b3-b872-9d5cb2e6cd04	Media archive	300046300
Abschlussarbeit ; Fotografie	eMuseum	300046300
Abschlussarbeit ; Fotomappe	eMuseum	300046300
Archivkarton Kontaktkopie	eMuseum	300046300
Archivkarton Kontaktkopie ; Abschlussarbeit	eMuseum	300046300
Ausstellung mit Dokumentarfotografie	eMuseum	300046300
Dokumentarfotografie	eMuseum	300046300
Foto Art	SIKART	300046300
Fotografie	eMuseum	300046300
Fotografie	SIKART	300046300
Fotografien	eMuseum	300046300
Fotosammlung	eMuseum	300046300
Photographie [Formschlagwort Sondersammlungen]	Graphic materials	300046300
Porträtfotografie	eMuseum	300046300
Sachfotografie	eMuseum	300046300
Werbefotografie für ein Parfum	eMuseum	300046300

## 6. Exploration of the Linked Data set

The sixth and last step concerned the exploration of the Linked Data. One objective of the LODZ project was to create a pilot application that proposes innovative search features using added semantic knowledge to explore the data. This step is based closely on the preceding ones. Due to some requirements on the data, the model and the transformation had to be reworked in order to output data that fits the needs of the application.

The result of this phase is a user interface, basically a search engine, that

1. offers a homogeneous access to the selected data resources on art and design in connection with the city and canton of Zurich,
2. provides flexible exploratory search options to the user by making use of the thesaurus relations in order to expand or limit the query,
3. allows a simplified user experience by presenting the added value of Linked Data with an aesthetically appealing visual design.

Particular attention was drawn to the usability of the application, balancing between the possible complexity of innovative search features and the required simplicity of current interfaces, due to the variety of access devices. Therefore, the usability was tested within the project team during a session dedicated to running a heuristic evaluation. Various wireframes were elaborated to conceive and preview the new features which were then implemented.

For the sake of a deeper technical understanding, some short explanations concerning the computer infrastructure are also given: the server implementation was realized using Java/Tomcat for the server side and Javascript/jQuery for the client side. The ad hoc developed Sesame based SQL/RDF mapper was realized based on MySQL<sup>(TM)</sup> and Blazegraph<sup>(TM)</sup>. The homogenized RDF datasets were integrated in a state-of-the-art triple store, hosted and operated on a Blazegraph<sup>(TM)</sup> repository by Semweb LLC.

The application itself is described in detail in the following section.

## Results

A tangible outcome of the LODZ project is a prototype application for art and design in Zurich, named ZHART (where "ZH" stand for Zurich). One particular objective was to develop a mash-up service on the application level to explore the added value of the resulting Linked Data. This was realized through four aspects visible on the interface as detailed below: the search entry points, the search engine results page, the thesaurus supported searching and the landing pages.

## Search entry points

On the entry page of the system the user can trigger a query by two different means (Figure 3): first, as common in all search engines, a simple search window can be used, by entering a free-text query. Alternatively, a tag cloud can be used to start a query with a given single-word-request. In an attractive layout, this tag cloud shows only the terms that are most frequently used within the four datasets. This enables the user to explore the content of the application instead of defining a premeditated query, thus allowing more serendipity.

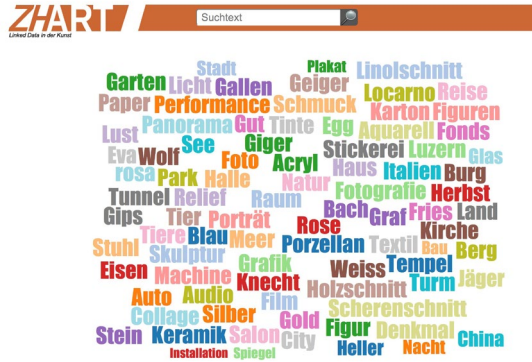


Figure 3. Homepage with two search entry points

Using the search field as well as clicking on one of the tags will lead to a search engine results page (SERP).

## Search engine results page

After having triggered a search action, the interface is divided into a left and a right pane (Figure 4). The left section shows the matches found in the thesauri regarding the main search terms (see next section). The right section shows a list of results found in the index built on top of the triple store.

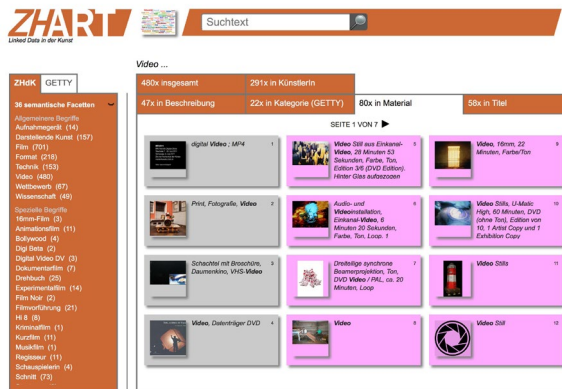


Figure 4. Search engine results page

The upper area of the right section shows a faceted navigation tab in which each facet indicates the exact distribution of the search term in the metadata fields. Those tabs, computed on the fly, can be used in the further search process, to refine the results list. As different datasets have been merged, the facets illustrate the data mapping with common RDF properties and contribute to more transparency for the search process.

The lower area of the right section shows a number of entities matched by the term(s) given in the search. As defined in the model (Figure 2), the two main entities of interest are work and person. Therefore, a result item is not necessarily a work, it can also represent a person. This is quite a novelty since in traditional library interfaces persons are not shown as such in the results. This change illustrates well how monolithic bibliographical records are broken up into different RDF concepts, allowing the retrieval engine to search rather on entities than on records. This progress is made possible through the RDF model and could be, in a future phase, extended to other concepts such as location or event.

In the SERP, further navigation can be done either by activating the pagination symbols (scrolling) or by clicking on one entity, showing an individual display.

### **Thesaurus-supported searching**

It is often said that RDF adds semantics to data. This added value was tested in another explicit search feature taking the form of a thesaurus-supported search functionality, as can be seen in the left section of the results page. Starting from the search query, it looks for matches in the GETTY AAT and eMuseum thesaurus, named simply GETTY and ZHDK (acronym of the Zurich University of the Arts) in the interface.

Once more, this feature was created to enhance the serendipity of the end user by proposing new or adjacent keywords for exploration, based on the experience made in "RODIN", a previous project of the Haute école de gestion de Genève (Belmonte et al. 2012). Consequently, the thesauri panel allows the user to explore related, narrower or broader terms. Each matched concept in one thesaurus is bound to further concepts which are shown together in a unique list for each of the relations "broader", "narrower" or "related".

### **Landing page**

Every entity in the data, be it a work or a person, has its own landing page that can be accessed via the results set. For a work, the landing page contains the following information: a thumbnail, a link to the picture in full resolution on the provider's website, the usual metadata for this work as well as a link to the author's landing page (if available). The landing page for the artist shows data describing the person's profile, e.g. his/her birth date, the field of activity and/or his/her nationality as well as a list of all his/her works available in ZHART (Figure 5).

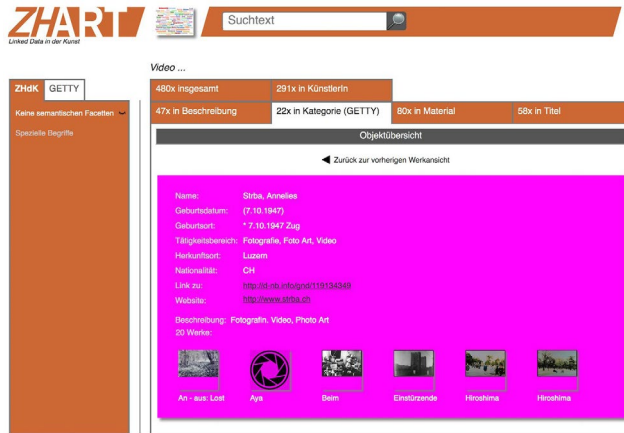


Figure 5. Landing page with the profile of an artist and his works

## Concluding discussion

The project Linked Open Data Zurich (LODZ) was realized on a small scale in order to gain experience with a concrete use case and existing data. The main goal was a) to present a proof of concept to integrate data with heterogeneous technical formats and b) to acquire skills to achieve this goal in the most efficient way. The creation of a stable and publicly available application would have required much greater resources, hence it was clear from the beginning that the outcome of LODZ should be considered as a prototype application and the basis for further decisions made by the project's principal stakeholders.

The four original datasets of the project come from very different institutions: a museum, a city and cantonal library, a university library and an art foundation. Besides that, the composition of the project team was quite diverse with partners from public, academic and commercial backgrounds. It was therefore not trivial to find a common language, especially for all questions concerning metadata management. This meant that the participants had to invest more time in communication and coordination aspects than expected. For that reason, the decision was taken to impose only a minimum of requirements for data delivery, although this subsequently required increased effort to standardise the heterogeneous data received.

This scale of diversity can be illustrated by the circa 30 different date formats that were found in the delivered data and the consequent effort required to match them. Without doubt, a more stringent control at the stage of the data delivery at the various institutions, conducted by information professionals, would have benefited the project, especially for some specific data fields such as the date. Prior decisions on this matter would also have allowed the creation of reusable automatic data validation tools.

Nevertheless, the most important challenge was the creation of links between the four datasets, and with external sources. This challenge has had some impact upon the resulting application, which offers fewer features with a semantic added value than expected. The interlinking possibilities were limited as the datasets were heterogeneous and had been created without prior consideration of common standards. To overcome these limitations, the choice was made to use a unique, generic thesaurus to create a bridge between the datasets and enable a single search entry point.

Another complementary approach to overcome this problem focused on manual interlinking, which required the collection of statistical information about the frequency of a particular field, and then about the half-normalised values within this field. This work was done on all four databases to enable a matching of the most frequent values towards a common reference dataset (in this case the GETTY AAT). Doing this manually is time-intensive and such a method cannot be taken into consideration for regular data updates; in this case, this work would have to be repeated regularly, which is not feasible. To prevent this inconvenience, the work of cataloguing in the original databases should focus more on the interoperability of data, for example by integrating external identifiers for persons or places rather than just terms in the form of strings. The potential of small Linked Data projects could thus be extended by such adaptations on the original data.

This may nevertheless imply, in some cases, new workflows of cataloguing and even new metadata management software products. As a consequence, the development of a Linked Data application would then be less time-consuming by leveraging the basic data quality, making it fit for interlinking, allowing instead a focus on adding new features, based on data enrichment and inferences.

The development of new search features is especially relevant in the field of art and design, where the creativity plays an important role. Do persons in this domain need a tool with exact and precise search functionalities, or rather a tool supporting inspiration? A hypothesis is that a particularly ludic interface, offering different or greater serendipity – possibly based on Linked Data – would benefit those working in a creative environment. This could be a subject for further research.

From a more pragmatic perspective, a better interoperability in the original data combined with the flexibility of the RDF model is a promising baseline for new applications that merge heterogeneous datasets. It enables work directly at the level of application with basis data, the latter being managed at a local level.

This advantage represents a great potential, especially for small datasets, the management of which cannot be delegated to a higher level and also lacks visibility, allowing the GLAM community to work more closely together towards a convergence of cultural metadata.

## References

- Belmonte, Javier, Eliane Blumer, Fabio Ricci, and René Schneider. 2012. "RODIN: An e-science tool for managing information in the web of documents and the web of knowledge." presented at the 3rd International Symposium on Information Management in a Changing World, Ankara, September 19. [http://by2012.bilgiyonetimi.net/proceedings/belmonte\\_blumer\\_ricci\\_schneider.pdf](http://by2012.bilgiyonetimi.net/proceedings/belmonte_blumer_ricci_schneider.pdf).
- Bensmann, Felix. 2016. "Swissbib data goes linked Teil 2: Verlinkung Und Anreicherung." *Swissbib Info*. April 28. <http://swissbib.blogspot.com/2016/04/swissbib-data-goes-linked-teil-2.html>.
- Bermès, Emmanuelle. 2013. "Convergence et interopérabilité: Vers le web de données." In *Le Web Sémantique En Bibliothèque*, Pt. 1, chap. 2. Collection Bibliothèques. Paris: Ed. du Cercle de la librairie. <http://www.electrelaboutique.com/ProduitECL.aspx?ean=9782765414179>.
- Berners-Lee, Tim. 2010. "Linked Data." *World Wide Web Consortium*. <http://www.w3.org/DesignIssues/LinkedData.html>.
- BNF. 2016. "About Data.bnf.fr." *Data.bnf.fr*. March 26. <http://data.bnf.fr/about>.
- Charles, Valentine. 2016. "Europeana data model documentation." *Europeana Professional*. <http://pro.europeana.eu/page/edm-documentation>.
- Coyle, Karen. 2016. "The technology." In *FRBR Before and After: A Look at Our Bibliographic Models*, 47–62. Chicago: American Library Association. <http://www.kcoyle.net/beforeAndAfter/978-0-8389-1364-2.pdf>.
- Dalbin, Sylvie, Emmanuelle Bermès, Antoine Isaac, Romain Wenz, Yann Nicolas, Tayeb Merabti, Anila Angjeli, et al. 2011. "Approches documentaires: priorité aux contenus." *Documentaliste-Sciences de l'information* 48 (4): 42–59. doi:10.3917/docsi.484.0042. <https://doi.org/10.3917/docsi.484.0042>
- Europeana. 2015. "Europeana linked open data." *Europeana Labs*. <http://labs.europeana.eu/api/linked-open-data-introduction>.
- HEG Genève. 2015. "Processus de transformation des données en RDF." *Filière Information Documentaire, Études Bilingues*. March. [http://campus.hesge.ch/id\\_bilingue/projekte/lodz/resultats\\_fr.asp](http://campus.hesge.ch/id_bilingue/projekte/lodz/resultats_fr.asp).
- Hügi, Jasmin, and Nicolas Prongué. 2014. "Les bibliothèques face aux Linked Open Data." Genève: Haute école de gestion de Genève. [http://doc.rero.ch/record/209598/files/M7-2014\\_memoire\\_HUGI-PRONGUE.pdf](http://doc.rero.ch/record/209598/files/M7-2014_memoire_HUGI-PRONGUE.pdf).

- Mäkelä, Eetu, Eero Hyvönen, and Tuukka Ruotsalo. 2012. "How to deal with massively heterogeneous cultural heritage data: Lessons learned in CultureSampo." *Semantic Web: Interoperability, Usability, Applicability* 3: 85–109.
- Schmachtenberg, Max, Christian Bizer, and Heiko Paulheim. 2014. "State of the LOD Cloud 2014." University of Mannheim, Data and Web Science Group. August 30. <http://linkeddatacatalog.dws.informatik.uni-mannheim.de/state/>.