

# Semantic BRICKS for Performing Arts Archives and Dissemination

*Sam Coppens, Erik Mannens, Rik Van de Walle (IBBT-MMLab-UGent)*

Until recently, cultural institutions in Flanders had little strategy to archive and disseminate their productions. And yet, the local government wants the productions to be archived as cultural heritage, schools want teaching packs for educational purposes, and other (foreign) institutions want production clips for promotional or research aims. Therefore, the following issues need to be tackled: a) the institutions want an easy-to-use, robust, decentralized archive; b) the institutions want to bundle and exchange their assets; c) the institutions want to use a common metadata schema combined with their own schemas; and d) the institutions want their (meta)data enriched and interlinked.

In this project, the problems investigated concern how the data should be archived and disseminated and what (meta)data schemas and application frameworks should be used. In this chapter, the main conclusions of this research are presented. Before jumping to conclusions, it is discussed which problems have to be tackled when digital material has to be archived and disseminated and how a semantic, layered (meta)data model answers these issues. In the following section, a semantic layered (meta)data model is developed that corresponds to the known problems of the previous sections. Following this, the relevance of a distributed framework that uses this semantic layered (meta)data model taking into account the special requirements of the cultural sector is shown. Furthermore, we elaborate on how assets can be bundled for dissemination. The next section

describes how data can be opened up and made available to the masses, whereas the following section shows how one can use the wisdom of the crowd by using the *Open Linked Data* principle the other way around. Finally, best practices and conclusions are drawn in the final sections.

## WHY GO LAYERED? A BRIEF LIST OF (META)DATA PROBLEMS

A digital repository offers numerous advantages besides archiving: spatial bordering blurs, mobility is no longer an obstacle, and searching through a vast number of files becomes much easier. But a digital repository is not invulnerable. Long-term preservation of digital multimedia data imposes specific requirements on the digital repository. First of all, the software and hardware of the digital repository must guarantee long-term access to the available information. Next, human intervention is still required both in the form of file descriptions, work processes and the use of standards to keep the information accessible and interpretable as long as possible for the user community.

Digital information is exposed to many threats. Some of these also endanger analogous documents, while others only target digital information:

- In digital form, information is a mere conceptual object. Digital multimedia can easily be copied and altered without immediate visible impact on the content therein. In comparison with analogue information, it is therefore difficult to guarantee the authenticity of digital information. Hence, one of the main concerns of long-term preservation is *permanently guaranteeing the authenticity* of data.

- Technological changes comprise another threat to digital data. Data formats and their inferred formats evolve rapidly. They can become obsolete or no longer interpretable in the future. The life span of storage techniques is also finite. To get rid of the discrepancy between the *short life span of digital technology* and the need for long-term preservation, either the old data format must be migrated to the new data format, or lasting emulation of the old data format must be foreseen. Moreover technical metadata must pass sufficient information onto the stored data to make fast interventions possible.
- In the long term, even the knowledge domain of the user community can change, data specialists come and go, or the institutions themselves can be modified or have a new task assigned. This possibly leads to interpretation problems. The stored data must therefore also contain *sufficient contextual metadata*, so new and future user groups can also still interpret the information.

When storing cultural heritage data from several different sectors digitally – i.e. broadcasters, libraries, the cultural sector or archival institutions – the digital repository will have to process a lot of descriptive metadata. Each specific scope stipulates which descriptive metadata are necessary. Digital images in a library can represent a scanned book, whereas images in the possession of a museum probably represent an artwork. Both images consequently demand other descriptive metadata fields. The digital repository must be able to search within these very divergent data sets. For this reason, a layered metadata system is necessary.

## WHY GO SEMANTIC?

### A BRIEF HISTORY OF KNOWLEDGE ON THE WEB

Nowadays, the web of hypertext is a fact. This web is actually a web of documents. These documents are described using Hypertext Mark Up Language (HTML). HTML is a language especially designed to describe web pages and the links between them. Such a web page usually consists of a body of text interspersed with multimedia objects, e.g., images, interactive forms, or movies. HTML provides a means to describe the structure of text-based information in a document. It is able to denote text as links, headings, tables, etc. This text is supplemented with embedded images, interactive forms and other objects. These HTML pages can be consulted using HTML browsers, e.g., Mozilla Firefox, which can present a web page in a human readable form.

```
<html>
  <body>
    <p>
      An image:
      
    </p>
    <p>
      Author: William Shakespeare
    </p>
  </body>
</html>
```

A lot of data presented by web pages comes from (relational) databases, spreadsheets, address books, etc. Unfortunately, HTML was not created to describe this kind of data. It can only describe a web page, which is a visual representation of that

data. To describe the data itself, eXtensible Markup Language (XML) was designed. XML is a set of rules for representing and structuring data in a textual format. Just like HTML uses tags and attributes to describe a web page, XML uses tags to describe a piece of data. XML parsers use these tags to extract the right piece of data from an XML document. XML documents are well-formed, because these XML documents are validated against an XML schema. An XML schema thus describes the structure of an XML document. This schema can, for instance, say that the value of the "author" tag must be "a string".

```
<document href="http://www.vti.be/examples/Hamlet.pdf">
  <author>William Shakespeare</author>
</document>
```

The next evolution was the eXtensible HyperText Markup Language (XHTML). XHTML restricted the rules of HTML to those of XML. It is actually a reformation of HTML to XML. This made it possible for XML parsers to parse XHTML documents, or to map XML documents automatically to XHTML representations. The benefits of XML-based web documents (i.e. XHTML) include searching, indexing and parsing as well as future-proofing the web.

XML has been the driving force behind the disclosure of a lot of (meta)data that is stored in databases, spreadsheets, technical drawings, etc. And yet, XML still has many interoperability issues. The same piece of information as in the previous example can be described in XML as:

```
<author>
  <uri>http://www.vti.be/examples/Hamlet.pdf</uri>
  <author>William Shakespeare</author>
</author>
```

Or as:

```
<document href="http://www.vti.be/examples/Hamlet.pdf"
author="William Shakespeare" />
```

These XML documents both describe the same piece of information, which is obvious for a human. For a machine parsing these two XML documents, these documents produce completely different XML trees. This makes it very difficult and syntax-dependent to query the XML tree. Furthermore, the tags used in the XML document do not mean anything to a machine. For a human, the tags already provide a hint of what their semantic meaning may be. This makes exchanging information using XML a significant task.

A solution is Resource Description Language (RDF). RDF describes information using triples. These triples consist of a subject ("http://www.vti.be/examples/Hamlet.pdf"), a predicate ("author"), and an object ("William Shakespeare"). Using these triples, any piece of information can be described by an RDF graph, which consists of a set of triples. These RDF graphs can also be described in a textual, interchangeable format, e.g., via RDF/XML, N3, Atom, etc. When these textual descriptions are parsed by a machine, they all end up with the same RDF tree. This is done by RDF reasoners, which build up the RDF tree. This makes querying the RDF tree, syntax-independent. Furthermore, all the nodes of the RDF tree are given a semantic meaning in RDF. For this purpose, RDF introduced namespaces. Namespaces are Uniform Resource Identifiers (URIs). By appending a namespace to the XML tags, those tags become unique, which makes it possible to define the semantics of that tag. A tag with a namespace thus forms the predicate in RDF. This allows information to be easily exchanged, re-using information and reasoning over that information.

An extension to RDF is RDF Schema (RDFS). RDFS is very similar to XML schema: it describes the structure of the RDF

document, and defines the semantics of its elements. It allows data to be structured with classes and properties on those classes. Another extension to RDFS is Web Ontology Language (OWL). OWL extends RDFS by introducing even more descriptive logic. For instance, it is possible to say that "all tragedies of William Shakespeare are plays", even if that current information is not included in the description of a specific play.

With these latest techniques, the *Semantic Web* is emerging. The machine-readable descriptions enable content managers to add meaning to the content, i.e. to describe the structure of the knowledge we have about that content. In this way, a machine can process knowledge itself using processes similar to human deductive reasoning and inference, thereby obtaining more meaningful results and helping computers to perform automated information gathering and research.

## A SEMANTIC LAYERED METADATA SCHEMA PROPOSAL

Metadata is actually data about data. Resources are fully described using metadata. It accompanies, for instance, a multimedia object, describing that multimedia object in a machine-readable way. This metadata is described by a metadata schema. These metadata schemas are very domain-specific, as every domain has different needs in describing their data. The major problem we are facing is to bridge the incompatibility of the different metadata schemas used all over the arts sector in Flanders (and beyond). Our proposed *layered ontology* will be used for the descriptive metadata in the project. This model not only leverages the exchange of data between the performing arts institutions in Flanders, but also the possible dissemination to the general public. The model has to be applicable in the whole performing arts sector in Flanders (and preferably beyond). In

other words, it has to be general enough. Many of the institutions already have descriptions of their objects. Those descriptions are formulated using many different metadata schemas. Therefore, it should be possible to map those schemas that are already in use in the performing arts sector in Flanders to our proposed layered schema.

The schema has to deliver all the necessary elements to the user so that he can find information on the object of his interest (i.e. general search). When the user has found his information, he has to be able to link to a more detailed description of that object (i.e. specific details). In order to fulfil these requirements the model is split into two parts, a description part (for the search) and a provenance part (for the detailed info).

The first part or common layer describes the object. This description has to be general enough to be applicable to all the objects in use, but on the other hand it has to deliver the elements so the user can find what he is searching for. This part consists of an interoperability layer, a common layer above all the metadata schemas that are already in use in the field. This part then automatically offers the tools to query all those descriptions. In other words it has to be able to answer basic questions like who, what, where and when (the famous 4 Ws).

The second part or lower layer contains the information needed to link to a more detailed description, mostly to the complete record the first part is mapped from. This part has to at least reflect the namespace of the schema the original record is described with, a URI of the repository the record comes from and the identifier of the record in that repository.

For the definition of the new metadata schema, we used W3C's Semantic Web technology, more specifically the OWL ontology language (as described in the previous section). The expressiveness of OWL allows fine-grained property definitions to be created by splitting the definition of properties into 'attributes'

and 'relations'. Attributes (corresponding to the OWL notion of a data type property) can take typed literals as value whereas relations (corresponding to the notion of an object property) can link to other resources such as content items or concepts taken from another ontology domain. The sublanguage is OWL DL, not OWL FULL. OWL FULL provides the most expressiveness, but does not guarantee the support of reasoning software, while OWL DL is a little less expressive, but it is guaranteed to be completely supported by the RDF reasoners. The framework BRICKS, which will make use of this schema and is described in the next section, also requires the schemas to be described in OWL DL.

The records are described in Dublin Core (DC). It is the most common metadata schema in use and it is general enough to describe all the objects of the Flemish performing arts sector. It is the largest common divider of all the metadata schemas that are used in the performing arts sector in Flanders. On top of that, all the fields of the DC model are optional and repeatable. This makes it possible to map nearly all the metadata schemas to DC. This also makes the Open Archives Initiative framework – Protocol Metadata Harvesting (OAI-PMH) compliant, because the offering of DC descriptions is a requirement for OAI-PMH compliance of the data provider. OAI-PMH is an XML protocol for harvesting metadata descriptions. It is used to *harvest and share* metadata. This protocol is a pillar within the BRICKS framework to import data from other OAI-PMH compliant repositories. For the implementation of the DC schema, all properties of DC were modelled as data type properties, which are all optional and repeatable.

As mentioned before, this lower layer should deliver at least three things: a) the metadata namespace of the originating record, b) the URI of the repository it comes from, and c) the identifier of that originating record in that repository. This layer is based on a schema that is used by the OAI-PMH protocol,

indicating the provenance of a record. This schema is described in an XML schema, so the schema was 'ontologised' in an OWL DL schema.

Finally, there needs to be an upper ontology that imports the two other ontologies and combines them into one ontology. This way each of the imported ontologies, the DC description (the common layer), and the Provenance description (the lower layer), can be altered independently.

## THE DISTRIBUTED OPEN-SOURCE BRICKS FRAMEWORK

After an initial platform evaluation the distributed semantic open-source repository BRICKS was chosen as a development platform. It is the outcome of the European project Building Resources for Integrated Cultural Knowledge Services (BRICKS). The aim of the BRICKS project was to design an open user- and service-oriented infrastructure to share structured knowledge and resources in the Cultural Heritage domain.

The key feature of BRICKS is its semantic, service-oriented, distributed architecture. The Service Oriented Architecture (SOA) of BRICKS means that its architecture is composed of several generic foundation components (called 'core' and 'basic' components). On top of this foundation layer a number of additional specialised services are implemented (called 'pillars'). Those services can be invoked by applications as remote services. These services are standard Web services described by WSDL documents. A BRICKS node (called 'BNode') is an application that uses these services. This allows BRICKS to be extended with other functionalities or services, and makes BRICKS an excellent development platform. There is already a basic BNode implementation, called 'Workspace', available to users, on which they can start developing.

The BRICKS architecture is *decentralised* by default and can be used out-of-the-box, thus every performing arts institution can deploy its own instance of BRICKS, called a BNode, without any problems. These BNodes are able to communicate with each other using P2P technologies. These BNodes can thus form a network of BNodes. This network allows, for instance, a search for data in all the BNodes within that network. Such an approach avoids having central hubs whose failure or overload could stop the whole system. Hence, BRICKS is a very heterogeneous, adaptable system without the need for a central body to maintain the system, making BRICKS a cost-effective solution, as centralised administration costs for additional personnel and money can be avoided.

The BRICKS framework uses OWL to describe its data. As discussed in the previous section, this semantic web technique describes your data in a very expressive, machine-readable way. This promotes data exchange, enriching your data with data from other datasets (as will be described in the following sections), and complex reasoning over your data. That is why BRICKS was chosen as an ideal initial development platform for the bulk of the Flemish cultural institutes.

## AGGREGATING RESOURCES THROUGH OAI-ORE

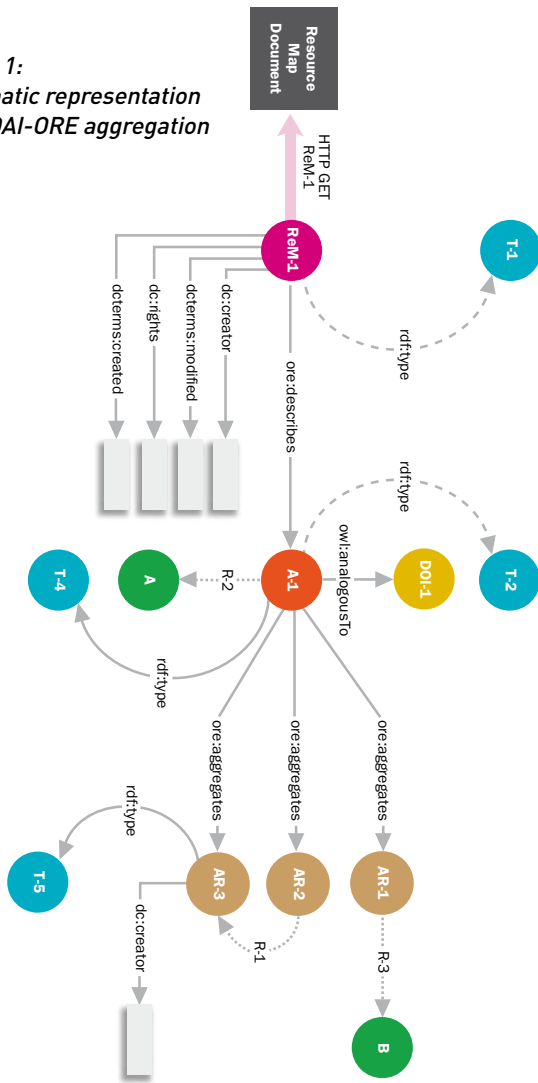
Besides archiving audio, video, photo, and text, the repository has to be able to store aggregations of these objects too. The performing arts institutions disseminate not only their performances but also introductions to performances, interviews with artists, programme brochures, reviews, etc. These aggregations have to be stored, disseminated, and exchanged too. For this, we developed an ontology based on the Open Archives Initiative Object Reuse and Exchange (OAI-ORE) protocol.

Today, many information systems, like content management systems, support the storage and identification of aggregations, and access to the aggregations and aggregated objects. In most systems, these objects vary in semantic type (e.g. article, book, video, dataset, etc.) and in metadata file format (e.g. PDF, XML, MP3, etc.). These objects can also be stored on different network locations, i.e., aggregated objects can be stored locally or externally. Information systems store, identify, and deliver access to these compound objects in an architecture-specific manner. Unfortunately, the way these information systems disseminate their compound objects is far from perfect and without any broadly accepted standard. In many cases, a lot of the advanced functionalities get lost when publishing the compound objects on the web. Mostly, the publication is aimed at the end-users (humans) and not at agents (machines) such as web crawlers. The structure of the object is often embedded in splash pages, user interface widgets, etc. This approach makes the structure of the compound object unclear for machine-based applications like browsers, web crawlers, etc. Consider the example of a scanned book, where all the pages get an HTTP URI. A web crawler can come across one of these pages and find links to the other pages of the book, to the chapter containing that page or to the book itself. A web crawler cannot distinguish between these links. For the web crawler these are untyped links or links that do contain information, but this information remains unreadable to the web crawler. Therefore, the order of the pages gets lost, etc.

The OAI-ORE standard tackles this problem by developing a standardised, interoperable and machine-readable mechanism that can express the information of compound objects. The standard makes sure that the logical boundaries of the aggregated objects and their mutual relations remain intact for machine agents when publishing the compound object on the web. To achieve this,

OAI-ORE makes use of *resource maps*. These resource maps are RDF (machine-readable) descriptions of the aggregation. They list the aggregated resources, their mutual relations and the web context of the aggregation, together with the URI of the resource it is describing, i.e., the aggregation. In fact, these resource maps are named graphs. These graphs are RDF graphs, sets of triples, extended with a name, a URI, for the graph/resource map. The named graph is not the aggregation itself, but a representation of its description encoded in Atom or RDF/XML, as depicted in Figure 1. The ORE model demands that a resource map describes just one aggregation. An aggregation, on the other hand, can have multiple resource maps, each with its own representation. This makes it possible to describe the same aggregation, for instance, with an RDF description and an XHTML description. Clients and applications need to determine the URI of the resource map from the URI of the aggregation, to get a description of the aggregation. This can happen in two ways: one way is to append a fragment identifier (“#”) to the URI of the resource map. For instance, the URI “http://example.com/aggregation” is the URI of the resource map, and “http://example.com/aggregation#” is the URI of the aggregation. In practice, this means that every aggregation should get a URI, just like any resource on the Web. From this URI, a web agent should be able to automatically get a machine-readable description of the aggregation, namely the resource map. Of course, this resource map also has a URI. This URI should be deducted from the URI of the aggregation. This is done, for instance, by using *cool URIs*. The web agent adds “.rdf” to the URI of the aggregation and gets its machine-readable description.

**Figure 1:**  
*Schematic representation  
of an OAI-ORE aggregation*



## LINKED OPEN DATA TO THE RESCUE

Sir Tim Berners-Lee first introduced the term Linked Open Data (LOD) in 2006. LOD lets people share structured data on the web as easily as they share documents today. It refers to a style of publishing and interlinking structured data on the web. LOD lets you use RDF data models to publish the structured data on the web and uses RDF links to interlink data from different datasets. This makes the web one giant database, the Web of Data.

LOD stipulates four basic principles. The first principle is that we first have to identify the items of interest in our domain. Those items are the resources that will be described in the data. The next principle is that those resources have to be identified by HTTP URIs (and avoid schemas such as Uniform Resource Names (URNs) and Digital Object Identifiers (DOIs)). The third principle is to provide useful information when accessing an HTTP URI. The fourth rule is to provide links to the outside world, i.e. to connect the data with data from other datasets in the Web of Data. This makes it possible to browse data from a certain server and receive information from another server. In other words, by linking the data with data from other datasets, the web becomes one huge database, called the Web of Data.

In practice, this means that every resource described by an RDF schema has to be identified by an HTTP URI, (e.g. "http://dbpedia.org/resource/Playwright"). Every resource should also have two representations: an XHTML (human readable) and an RDF (machine-readable) representation. Every representation also has to be identified by an HTTP URI (e.g. "http://dbpedia.org/page/Playwright") for the XHTML representation and for the RDF representation (e.g. "http://dbpedia.org/data/Playwright"). When coming across the HTTP URI of a resource, the LOD server determines which representation should be served, based on information in the Accept header of the user's client, and redirects



the client to the appropriate representation using HTTP's 303 redirect and content negotiation.

Publishing resources as LOD, conforms to the way OAI-ORE offers to publish aggregations. OAI-ORE demands that aggregations have to be identified by a URI, and have to be described using an RDF schema, i.e. a resource map, which also has a URI. When clients use the URI of that aggregation, they should be able to automatically detect the URI of the resource map with the appropriate representation for the client. This principle conforms to the way LOD publishes data, except that with LOD, the client gets automatically redirected to the appropriate representation, based on the client Accept header (which is a benefit).

For publishing the records from a triple store as LOD, the open-source tool Pubby was used. Pubby is actually a Linked Data frontend for SPARQL endpoints. A SPARQL endpoint is a web service that can handle SPARQL queries. These SPARQL queries can be seen as semantic SQL statements. BRICKS does not provide such a SPARQL endpoint. That is why the triple store in the BRICKS framework was replaced by the open-source OpenLink Virtuoso triple store. This triple store offers a SPARQL endpoint by default. By configuring Pubby for the SPARQL endpoint, provided by the Virtuoso triple store, the records stored in the triple store are published as LOD. This means providing HTTP URIs for all the records served by the SPARQL endpoint, providing a simple HTML interface showing the data available about each resource, and taking care of the automatic redirecting to the appropriate representation.

In fact the BRICKS framework has no problems storing the resource maps, but cannot handle the cool URIs. Within BRICKS you cannot define your own URIs. This problem is solved by publishing the records from the JENA triple store from BRICKS as LOD, as was described above. This way, you get full control over the URIs used. Publishing the records as LOD

offers the opportunity to use cool URIs to redirect the client (web crawlers, HTTP browsers, machine agents) to the appropriate representation. This way, clients that come across the HTTP URI of an aggregation can be redirected to the resource map, a representation they understand, preserving the typed links between the aggregated resources. Hence, storing the resource maps and publishing the resource maps as linked data makes the repository OAI-ORE compliant. This allows the BRICKS repository to manage, exchange, and share aggregates of resources, e.g. a video of a performance, accompanied by a program brochure and a transcription of the performance, conforming to the OAI-ORE standard. Because the records are published as LOD, the publishing of the records is not handled by the BRICKS platform anymore. It becomes solely an administration platform, regulating the imports into the triple store.

## GETTING EVEN MORE THROUGH METADATA ENRICHMENT

Finally, the stored records, constructed via our metadata schema and published as LOD, are extended with links to information from datasets like GeoNames and DBpedia. This way, the records are enriched with information from external datasets, weaving that extra information into the Web of Data.

To enrich the data automatically, the choice was made to provide extra information on the title of the resource, the people, organisations, events, and the places involved. In practice, this means iterating all DC descriptions of the records and seeing if there are people, organisations or events in its DC descriptions. When such concepts exist in the description, the DBpedia dataset is queried, asking for information about that concept. The same is done for places, but for these concepts the GeoNames dataset is queried. The results, returned from these queries, are HTTP URIs

with extra information on the requested topic. This HTTP URI is added to the DC description via the object property: `rdfs:seeAlso`.

The descriptions of the resource (values from the `dc:description` datatype property from the DC description) are also examined. These strings, describing the resource, are investigated for people, organisations, companies, brands, locations, and events. For this, we rely on the OpenCalais web service, which is able to investigate strings and return certain concepts mentioned in the description. The results for the people, organisations, or events concept are forwarded to query the DBpedia dataset. The results for the places concepts are forwarded to query the GeoNames dataset.

By applying our metadata enrichment algorithms, the records are enriched with links to information from an external dataset. This not only puts the records on the Web of Data, but also enriches these records with extra information.

## BEST PRACTICE SUMMARY

The solution proposed in this chapter elaborates on the distributed semantic open-source BRICKS archiving and distribution architecture, since ease of use, robustness, independence of central authorities, low cost, and flexibility in offered services are crucial within the cultural community. This platform allows the institutions to configure, extend and manage their own digital repository according to their needs. In order to store and exchange all the information on their productions, a new layered metadata schema is developed on top of the BRICKS framework. This is an OWL DL schema consisting of two layers: Dublin Core and Provenance. The Dublin Core layer describes the digital objects in a general way as a greatest common divisor. All the fields of Dublin Core are optional and repeatable. These characteristics allow for easy mapping to and the adoption of the

proposed metadata schema. It forms a common interoperability and discovery layer on top of the descriptions that are already distributed by the institutions. The second layer indicates the provenance of the Dublin Core descriptions. In most cases, the institutions have their own metadata schema which is mapped to Dublin Core. The provenance layer indicates the identifier of the original metadata description and the namespace of the original metadata schema. This information allows linking to the original descriptions, which are in most cases richer in information. To aggregate the digital objects in bundles (for educational purposes among other things) the BRICKS framework is extended with an OAI-ORE web service. It describes aggregations of Web resources in a semantic way via dereferencable URI's. Furthermore, we enrich the metadata semantically following the Linked Open Data principle. In our case, we apply linguistic processing on the plain text contained to various elements of the metadata such as title, contributor, subject, and description. The linguistic processing consists in extracting named entities such as people, organisations, companies, brands, locations, and events using the OpenCalais infrastructure. Once the named entities have been extracted, we map them to formalised knowledge on the web available in GeoNames, for the locations, or in DBpedia, for the people, organisations, and events, and feed this new knowledge back into the system. This way, BRICKS is semantically adapted and extended to offer an end-to-end solution to the institutions and third parties (schools, broadcasters, etc.) that can search, harvest, and publish all data via web services.

## CONCLUSION

This chapter showed how performing arts institutions can disseminate their content using semantic web technologies, like RDF, OWL, and Linked Open Data. The Semantic Web is an evolving extension of the World Wide Web in which the semantics of information and services on the web is defined, making it possible for the web to understand and satisfy both the requests of people and machines to use the web content. To benefit the search and discovery of the records, these records have to be described by a *uniform metadata model*. This model has to be applicable for a variety of data: text, audio, video, and aggregations of them. For this purpose, three semantic models were designed and implemented: a Dublin Core description, describing the resource in a very generic way, a provenance description, referencing the original record, which can give a more detailed description of the resource than the Dublin Core description, and an OAI-ORE model to describe aggregations. This way, the performing arts institutions can share and exchange their (aggregations of) information, avoiding many interoperability issues. By publishing the records in a Linked Open Data way, the server can redirect clients (people or machines) to the appropriate representation, XHTML for people and RDF for machines, which is compliant to the way OAI-ORE publishes aggregations. By further enriching the data with links to information coming from DBpedia and GeoNames for instance, the more expressive records are weaved into the Web of Data, making the Web of Data one huge database.

As such, we showed how all performing arts productions media can be archived, bundled and disseminated using distributed Semantic Web technologies. In the end, everything is demonstrated within an end-to-end Proof-Of-Concept showing the feasibility of the approach in Flanders' cultural institutions,

establishing a durable cooperation between all actors involved where a) the institutions have an easy-to-use, robust, decentralized archive; b) the institutions can bundle and exchange their assets; c) the institutions can use a common metadata schema combined with their own schemas; and d) the institutions have their (meta) data enriched and interlinked.

This is a new approach for disseminating records from the performing arts sector. Mobilising the sector to adapt this approach is not a trivial task, although the awareness comes from the sector itself. This is why VTi, the Flemish Theatre Institute, as a coordinating body for the performing arts in Flanders, chose to implement this approach first of all and to offer this approach as a service to the other institutions in the performing arts. This way, the institutions are more easily mobilised and encouraged to adopt this way of disseminating archived multimedia of the performing arts produced in Flanders.