Project Acronym: INDICATE

Project Title: International Network for a Digital Cultural Heritage e-Infrastructure

Contract Number: 261324

Starting date: 1 September 2010 **Ending date:** 31 August 2012

Deliverable Number: D 4.1

Title of the Deliverable: Pilot on semantic search

Task/WP related to the Deliverable: Task 4.1

Type (Internal or Restricted or Public): Public

Author(s): Nikolaos Simou, Costas Pardalis

Partner(s) Contributing:

Contractual Date of Delivery to the CEC: Month 18

Actual Date of Delivery to the CEC: March 2012

Project Co-ordinator

Company name : Istituto Centrale per il Catalogo Unico (ICCU)
Name of representative : Rosa Caffo
Address : Viale Castro Pretorio 105, I-00185 Roma
Phone number : +39.06.49210427
Fax number : +39.06. 06 4959302
E-mail : rcaffo@beniculturali.it
Project WEB site address : http://www.dc-net.org

Context

WP 4	Research Pilots
WPLeader	NTUA
Task 4.1	E-Infrastructures-enabled semantic search
Task Leader	NTUA
Dependencies	
Starting date	Month 1
Release date	Month 18

Author(s)	Nikolaos Simou, Kostas Pardalis
Contributor(s)	Tasos Venetis, Despoina Trivela
Reviewers	Sergi Fernández, Mercè López

HISTORY

Version	Date	Author	Comments
01	17-02-2012	Nikolaos Simou	First version, index and structure
02	25-02-2012	Tasos Venetis	First Draft of section "Deployment of Semantic Search Pilot on e-Infrastructures"
03	25-02-2012	Despoina Trivela	First draft of section "Semantic Representation of MICHAEL content"
04	05-03-2012	Kostas Pardalis	Revision of section "Deployment of Semantic Search Pilot on e-Infrastructures"
05	07-03-2012	Nikolaos Simou	Executive summary and conclusions are added. Overall check of document
06	21-03-2012	Sergi Fernández, Mercè López	Internal review
1.0	28/3/2012	A. Fresa	Final version submitted to EC

Table of Contents

1 Executive summary

This deliverable describes the pilot of the semantic search using e-infrastructures. For the implementation of this pilot we have used the MICHAEL dataset. Firstly, the metadata were transformed to RDF providing in that way a semantically richer representation than XML, that was the initial format. During this process, which is also known as RDFization, the elements of the XML documents were mapped to RDF classes and properties that were used for the semantic representation. After that, the values of specific elements of the dataset, like countries, persons and languages were used for the discovery of resources that describe them in external sources. The discovered resources serve additional information for the person, country, and language of interest than the information originally provided in the dataset. This extra information added through this procedure is very useful for increasing the searchability of the content. Therefore, through this process, semantic enrichment of the MICHAEL dataset was established. The transformed to RDF and enriched metadata that resulted from the previous modules were then stored in a semantic repository which was semantically accessible by using the SPARQL query language. The most important outcome of this procedure was the support of queries for the retrieval of MICHAEL content that were composed of concepts and properties using information that was not available on the original dataset.

The aforementioned workflow was deployed on Amazon Elastic Compute Cloud (EC2), considered as one of the most mature Cloud e-infrastructures at the moment. For the deployment on the Amazon Elastic Compute Cloud a processing interface (MINT-PI) was implemented that parallelizes the semantic transformation and enrichment taking in that way advantage of the massive processing power that can be offered by a Cloud infrastructure. Furthermorem, for further utilising the processing power of e-infrastactures and also for increasing the overall scalability of the architecture, the semantic repository used for storing the semantically enhanced dataset was also deployed on EC2.

Evaluation was performed for the enrichment process, as well as for the overall performance of the proposed architecture. More specifically, the algorithm developed for the resource discovery operated outstandingly well for the countries and the languages, while very good results were also established for persons considering the difficulty of the task. Finally, comparing the overall performance of the workflow when deployed on cloud e-infrastructures and when deployed on a single server it was observed that e-infrastructures can provide scalability that is an essential characteristic when processing massive volume of content –that is always the case for digital cultural heritage - with very low cost.

2 Introduction

In the last few years, Digital Cultural Heritage Domain has known great evolution. Massive digitisation and annotation activities have been taking place all over Europe and the United States following the early developments at European level and the Lund principles [1]. The increasing support of European Union and the positive reaction of companies like Google, who have shown great interest and are strongly involved in the domain, have led to a variety of, rather converging, actions towards multimodal and multimedia cultural content generation from all possible sources (i.e. galleries, libraries, archives, museums, audiovisual archives, etc.).

The main difficulty arising when we attempt to manipulate cultural content is related to the diversity of types and schemas used to describe that content. The use of different formats while expressing cultural information discourages any kind of automatic process related to search and retrieval. Semantic interoperability is considered to be the solution to overcome such difficulties [2],[3]. Indeed, the representation of exchanged information using common and well-defined semantics enables the automatic interpretation by systems. Semantic interoperability is achieved by using representation languages that define the formal semantics of information and therefore, establish a certain level of common understanding that renders information machine-understandable. Once systems understand data representation semantics, they are able to process their content and provide searching and retrieval services. Semantic Web languages widely used for this purpose are Resource Description Framework (RDF) [4] and Web Ontology Language (OWL) [5].

In this work we established a semantic search system using MICHAEL dataset, the fruit of MICHAEL project [6]. Furthermore, we evaluated the feasibility of semantic search while using e-Infrastructures and analysed the benefits gained from this integration. The rest of this document is organised as follows. The first section highlights the overall system architecture. The next two sections describe in more detail the data manipulation process, that is, the transformation of data to RDF and their semantic enrichment. Section 5, illustrates the deployment of semantic search system on e-Infrastructures while in Section 6 the evaluation of the proposed architecture is presented. Finally, in the last section a discussion on the benefits of using e-Infrastructures for the semantic enhancement and retrieval of cultural metadata is given.

3 Overall Architecture

The main objective of the pilot was to provide an e-Infrastructure enabled semantic search service for cultural repositories. The overall architecture is demonstrated in the figure below.

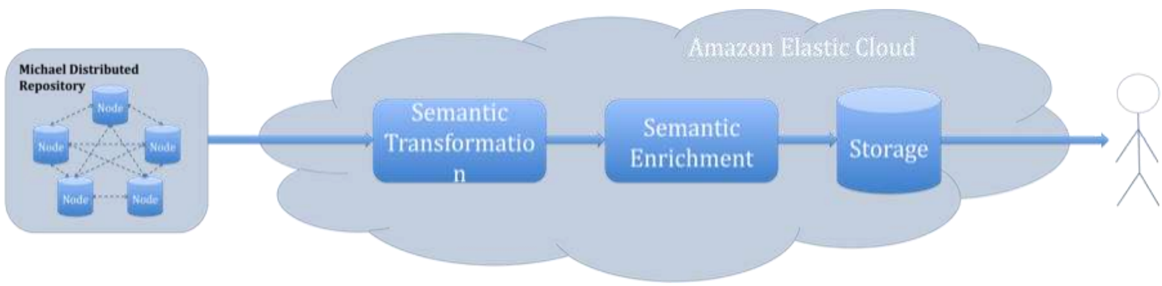

Figure 3-1 Overall architecture

The metadata used for the pilot came from the MICHAEL project and are stored in several geographically distributed XML based repositories. Therefore, the first step is to harvest the metadata from these repositories, a process that is performed by using the data delivery protocol OAI-PMH [8].

The next step is the semantic transformation during which the MICHAEL metadata are transformed from XML to RDF, a semantically richer representation that enables machine-understandable semantics. After this process, that is also called RDFization, the semantic enrichment of the metadata can be performed. Once metadata is in RDF it is possible to identify instances and map them to external resources. Indeed, specific values of the examined dataset are discovered as DBpedia resources [9]. Linking to DBpedia resources provides additional semantic information related to them and thanks to this data enrichment, more expressive queries can be posed over the semantic repository. The outcome of those processes, which produces a common understanding of the content for machine agents, is stored to an RDF repository. The end user can then semantically search for the MICHAEL content using the SPARQL query language [7].

The aforementioned modules are deployed on Amazon Elastic Compute Cloud (EC2). By this deployment we managed to examine the technical requirements for semantic search on e-infrastructures and also to outline the advantages and disadvantages of this architecture.

4 Semantic Representation of MICHAEL content

4.1 MICHAEL Dataset

The MICHAEL project was funded through the European Commission’s e-Ten programme, to establish a new service for the European cultural heritage. The main idea behind the project is the integration of national initiatives in digitisation of the cultural heritage and interoperability between national cultural portals to promote access to digital contents from museums, libraries and archives. It has established an international online service to allow users to search, browse and examine descriptions of resources held in institutions from across Europe. MICHAEL data form multilingual digital cultural heritage inventories available to all through an open source technical platform.

MICHAEL data were therefore used to run the INDICATE pilot in order to demonstrate the value of the semantic enrichment and to learn lessons about the deployment on the cloud computing.

4.2 Transformation of MICHAEL dataset to RDF

The transformation of MICHAEL metadata to RDF is necessary for establishing machine understandable semantic content. Data values are converted to resources that are then, linked to external data sources like DBpedia whenever possible. The transformation is not trivial due to the different purpose that each format is intended to serve. XML is used for collecting metadata about cultural content, while RDF is employed to make statements about Web resources in the form of subject-predicate-object expressions, so called triples. Therefore, during the RDFization the things described in the XML document have to be firstly identified, together with the statements about them, before proceeding with the instantiation of the RDF document.

Although RDF provides a generic, abstract data model for describing resources using subject-predicate-object triples, it does not provide any domain-specific terms for describing classes of things in the world and how they relate to each other. This function is served by taxonomies, vocabularies and ontologies expressed in SKOS (Simple Knowledge Organization System) [15], RDF, RDFS (the RDF Vocabulary Description Language, also known as RDF Schema) [16] and OWL (the Web Ontology Language). Hence, a decision that was made in accordance with the objects described in the MICHAEL XML documents was the selection of the vocabularies used for the RDF representation. The most appropriate vocabularies for representing the content of MICHAEL was FOAF [17], relationship [18], DCMI Type vocabulary [19], several schemas, DC [20], DCTerms [21], RDF and RDFs, as well as some ontologies, Taxonconcept ontology [22], the Ordered List Ontology [23] and the Counter Ontology [24].

Once the selection of the appropriate vocabularies for the RDF representation is complete, we are ready to create the resources for the things described in the xml document. Thus, the described objects are matched to a URI. In detail, for every XML document an RDF file was constructed and named after the unique identifier ID, found in the XML accompanied by RDF file extension. The same ID was then used in order to construct the resource, identifying each item served under the domain: http://mint.image.ece.ntua.gr/. For example, an item assigned with the unique identifier 'UK-DC-7036a5f' is mapped to the resource http://mint.imae.ece.ntua.gr/resource/UK-DC-7036a5f in the UK-DC-7036a5f.rdf file.

The next step is to map XML elements to RDF properties using the selected vocabularies. For example, XML element 'period' is mapped to 'dc:date' and element 'famousEvent' to Counter ontology's property 'co:event'. In the following table we present the mappings of the MICHAEL metadata XML elements to RDF properties.

id	DC_11:identifier	Subject	DC:subject
Title	DC_11:title	Period	DC:date
Description	DC_11:description	SpatialCoverage	DCTerms:spatial
StartDate	Taxon:startDate	Culture	DC:type
EndDate	Taxon:endDate	FamousPeople	Co:event
Rights	DCTerms:rights	FamousEvent	Olo:item
Language	DCTerms:language	FamousItem	Dcterms:coverage
DigFormat	DCTerms:format		

Table 1 Mappings of XML elements to RDF properties

Another aspect that is very important when converting metadata to resources is the provenance of information. Therefore, for every RDF representation of an item, provenance metadata has been published that include the publication date and the creator, allowing in that way consumers to track the origin of particular data fragments. (The interested reader in the exact XML to RDF transformation is referred to appendix where an XML file and its corresponding RDF representation are given.)

4.3 Enrichment

As far as enrichment techniques are concerned, the most trivial one is manual enrichment, where items are manually annotated with new fields. Workflow in data processing may be facilitated when some values for new fields are suggested automatically by appropriate algorithms. In this semi-automatic technique the final value selection depends again on the user. The applied enrichment technique is an automatic one based on text mining tools. Specific values served in the examined dataset such as names of countries, persons and languages, are compared with the values of resources served by DBpedia using SPARQL and, if a match is found, a link is created. Thus, country England is linked to http://dbpedia.org/resource/England and so on. Then, it can be used as the object of a triple that has as predicate the DCTerms property "spatial" and as subject the item resource:

<subject>	*<property>*	*<object>*
http://mint.image.ece.ntua.gr/resource/UK-DC-7036a5f	DcTerms:spatial	*http://dbpedia.org/resource/England*

Similarly, person's names appearing as value of several elements in several items are identified by the corresponding DBpedia resource. Thus, John Hawkins is linked to http://dbpedia.org/resource/John_Hawkins and forms the object of relationship ontology property 'participant' which maps to element 'famousPeople':

<subject>	*<property>*	*<object>*
http://mint.image.ece.ntua.gr/resource/UK-DC-7036a5f	Relationship:participant	*http://dbpedia.org/resource/John_* *http://dbpedia.org/resource/John_Hawkins*

The outcome of this process is very important for the retrieval of the content since the linking to an external source brings to our disposal all additional information served by the source. In the following table we illustrate some evaluation results concerning data enrichment. The first column indicates the number of instances found in Michael dataset and the second one the number of DBpedia resources that were mapped to the corresponding instance. We observed a significantly high percentage of countries and language instances included in Michael data that was linked to DBpedia.

	Total	Found	Percentage
Countries	16429	15987	97.3%
Languages	11090	11032	99.5%
Persons	6442	3632	56.4%

Table 2 Evaluation results

As explained above, semantic interconnections of content descriptions with external ones published on the web serve to provide the ability to pose expressive queries. The services of publishing and querying are heavily depended on how the data are enriched and stored. The transformed MICHAEL dataset that is linked to external sources should be stored in a repository that facilitates the services of search, retrieval and publishing. As it is described in Section 5.3, when e-Infrastructure technologies are employed, the delivered services are greatly enhanced since they provide scalability, vital for semantic enrichment and querying.

5 Deployment of Semantic Search Pilot on e-Infrastructures

The complexity of performing the aforementioned procedures is computationally expensive due to the vast volume of content that is always the case for digital cultural repositories. Furthermore, the enrichment process requires checks using the network connection for the discovery of external resources, fact that makes it very expensive. If we also consider enrichment in datasets like the MICHAEL, that contain more than one value that can be discovered as resources in external sources, then it is easy to understand that hours of processing are required. For that reasons the use of e-Infrastructure for the management of digital cultural content is essential.

5.1 Overview of e-Infrastructures

An infrastructure can be generally defined as the set of interconnected structural elements that provide framework supporting an entire structure of development. Nowadays, the evolution of Internet has given rise to e-Infrastructures,technologies that support collaborations built upon an infrastructure of grid computing software that can provide benefits such as shared access to large data collections, advanced tools for data analysis and large-scale computing. They embrace networks, grids, data centres and collaborative environments, and can include supporting operations centres, service registries, single sign-on, certificate authorities, training and help-desk services. Most importantly, it is the integration of these that defines e-Infrastructure. Two of the most important elements of e-Infrastructure are the distributed computing infrastructures: grid and cloud computing.

5.1.1 Grid computing

Grid computing is a term referring to the federation of computer resources from multiple administrative domains to reach a common goal. The grid can be thought of as a distributed system with non-interactive workloads that involve a large number of files. What distinguishes grid computing from conventional high performance computing systems such as cluster computing is that grids tend to be more loosely coupled, heterogeneous, and geographically dispersed. Although a grid can be dedicated to a specialized application, it is more common that a single grid will be used for a variety of different purposes. Grids are often constructed with the aid of general-purpose grid software libraries known as middleware.

Grid size can vary by a considerable amount. Grids are a form of distributed computing whereby a 'super virtual computer' is composed of many networked loosely coupled computers acting together to perform very large tasks. For certain applications, 'distributed' or 'grid' computing, can be seen as a special type of parallel computing that relies on complete computers (with onboard CPUs, storage, power supplies, network interfaces, etc.) connected to a network (private, public or the Internet) by a conventional network interface, such as Ethernet. This is in contrast to the traditional notion of a supercomputer, which has many processors connected by a local high-speed computer bus.

5.1.2 Cloud computing

Cloud computing is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices as a metered service over a network (typically the Internet). Cloud computing provides computation, software, data access, and storage resources without requiring cloud users to know the location and other details of the computing infrastructure.

End users access cloud based applications through a web browser, a light weight desktop or a mobile app, while the business software and data are stored on servers at a remote location. Cloud application providers strive to give the same or better service and performance as if the software programs were installed locally on end-user computers.

At the foundation of cloud computing is the broader concept of infrastructure convergence (or Converged Infrastructure) and shared services. This type of data centre environment allows enterprises to get their applications up and running faster, with easier manageability and less maintenance, and enables IT to more rapidly adjust IT resources (such as servers, storage, and networking) to meet fluctuating and unpredictable business demand.

5.2 Amazon Elastic Compute Cloud

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable computing capacity in the cloud. It is one of the most technologically mature Cloud environments with a simple web service interface that allows obtaining and configuring capacity with minimal friction. It provides complete control of computing resources and allows running on Amazon's proven computing environment. Additionally, it is used as the initial Cloud environment for deployment and further benchmarking as it provides a concrete pricing model for comparisons. It moreover reduces the time required to obtain and boot new server instances to minutes, allowing to quickly scale capacity, both up and down, as computing requirements change. Finally it provides developers the tools to build failure resilient applications and isolate themselves from common failure scenarios.

5.2.1 Amazon EC2 services

In the following paragraphs we present an overview of the highlights of Amazon EC2 services.

First of all, Amazon EC2 is *elastic*, as it enables users to increase or decrease capacity within minutes. Users can commission one, hundreds or even thousands of server instances simultaneously and as this is all controlled with web service APIs, each application can automatically scale itself up and down depending on its needs.

Moreover Amazon EC2 is *completely controlled* as users have full control of their instances. Root access is provided for accessing each one, making interaction with them as simple as interacting with any machine. More interestingly any instance can be stopped while retaining data from the boot partition and then can be restarted using the web service APIs.

One of the most important services that Amazon EC2 provides is *flexibility*. It gives the users the choice of multiple instance types, operating systems, and software packages. It even allows the selection of a user-specified configuration of memory, CPU, instance storage and the boot partition size that is optimal for every specific choice of operating system and application.

Additionally, Amazon EC2 works in conjunction with Amazon Simple Storage Service (Amazon S3), Amazon Relational Database Service (Amazon RDS), Amazon SimpleDB and Amazon Simple Queue Service (Amazon SQS) to provide a complete solution for computing, query processing and storage across a wide range of applications.

Amazon EC2 provides numerous mechanisms for securing compute resources:

- It includes web service interfaces to configure firewall settings that control network access to and between groups of instances.
- When launching Amazon EC2 resources within Amazon Virtual Private Cloud (Amazon VPC), compute instances can be isolated by specifying the IP range. This way, users can connect to the existing IT infrastructure using industry-standard encrypted IPsec VPN. Dedicated Instances, that are Amazon EC2 Instances that run on hardware dedicated to a single customer for additional isolation, can also be launched into VPC.

5.2.2 Indicate cluster services

The Indicate Cluster consists of a large instance of 7.5 GB of memory, 4 EC2 Compute Units (2 virtual cores with 2 EC2 Compute Units each), 850 GB of local instance storage in a 64-bit platform. Moreover it utilises Elastic IP addresses assigned to each instance in order to ensure the existence of static IPs. Finally Amazon Elastic Block Store (EBS) has been used in order to provide persistence storage to the Indicate Cluster Instances.

5.3 Semantic Repository for data storage

In order to store the data produced by the RDFization process we used databases that were purpose-built for the storage and retrieval of RDF metadata. These databases are called triplestores and they were set up on the provided infrastructure. In the following we present some of the most popular triplestores that support grid computing.

Bigdata [26] is a clustered RDF store for ordered data (B+Trees) and designed to run as a server on commodity hardware. It is available under the GNU General Public License (GPL) and is designed for UNIX systems (Linux) with client connectors available for Java (Sesame). Additional scale can be achieved by simply plugging in more data services dynamically at runtime, which will self-register with the centralized service manager and start managing data automatically. Scale-out is achieved via dynamic key-range partitioning of the B+Tree indices.

OWLIM [12] is a family of commercial RDF storage solutions, provided by Ontotext. It is available in two different editions: SwiftOWLIM is designed for medium data volumes (up to 100 million triples) since reasoning and query evaluation are performed in main memory, while BigOWLIM is designed for large data volumes and uses le-based indices that allow it to scale up to billions of RDF triples supporting dynamic configuration of cluster. Additionally, OWLIM is available as a SAIL (Storage and Inference Layer) for the Sesame RDF framework. SwiftOWLIM source code is provided free of charge for any purpose under a GNU LGPL license. BigOWLIM is free to use for research, evaluation, and testing purposes; for commercial applications an appropriate license is required. Both versions do not provide a dedicated extensibility mechanism but allow the definition of custom rules and rule languages for the inferencing process. BigOWLIM is in use for a large number of Semantic Web and Linked Data applications.

4Store [10] is an RDF database developed by Garlik Inc. It is implemented in ANSIC99 and available under the GNU General Public License (GPL), version 3. It is designed for UNIX-like systems (Linux, Mac OS) and runs as a server on a single-machine or in cluster-mode on 64bit machines. Client connectors are available for PHP, Ruby, Python, and Java. Dedicated extensibility mechanisms are not foreseen.

SHARD [27] which stands for Scalable, High-Performance, Robust and Distributed is an open source cloud-based triplestore technology that enables scalable data processing and analysis based on the Cloudera Distribution for Hadoop implementation of the MapReduce formalism. SHARD triple-store persists data as standard RDF triples and runs queries over this data using the standard SPARQL query-language.

Dydra [28] is a cloud-based RDF store, a database-as-a-service. It is a multi-tenant data store and query engine. The user does not have to estimate the data size, worry about clusters, nodes, resource use, or make big licensing commitments up front.

Sesame [11] is an open source Java framework for storing, querying and reasoning with RDF and RDFS. It can be used as a database for RDF and RDF Schema, or as a Java library for applications that need to work with RDF internally.

In order to select a triplestore for the storing of the data in Indicate the following requirements had to be met:

- The triplestore had to be *distributed*. This means that the storage devices are not connected to the same CPU, but they may be stored in multiple computers located in the same physical location, or may be dispersed over a network of interconnected computers.
- The triplestore had to have an open source license. An open-source license is a copyright license for computer software that makes the source code available for everyone to use. This allows end users to review and modify the source code for their own customization and/or troubleshooting needs. Open-source licenses are also commonly free, allowing for modification, redistribution, and commercial use without having to pay the original author.
- The triplestore had to support SPARQL that is an RDF query language designed to meet the use cases and requirements identified by the RDF Data Access Working Group.
- Finally the triplestore had to be Web accessible.

5.4 Processing Infrastructure

For the deployment of the data manipulation services on the cloud their parallelization was essential. In order to achieve this objective a processing interface based on messaging was implemented. In this section we shortly present the basic messaging techniques, the RabbitMQ on which our PI is based and MINT PI that was implemented for the pilot parallelization.

5.4.1 Messaging

Messaging is a form of loosely coupled distributed communication, where in this context the term 'communication' can be understood as an exchange of messages between software components. Message-oriented technologies attempt to relax tightly coupled communication (such as TCP network sockets, CORBA or RMI) by the introduction of an intermediary component. This approach allows software components to communicate 'indirectly' with each other. Benefits of this include message senders not needing to have precise knowledge of their receivers.

The advantages of messaging include the ability to integrate heterogeneous platforms, reduce system bottlenecks, increase scalability, and respond more quickly to change. In the following we present the overview of some of the current state of the art messaging protocols.

5.4.1.1 AMQP

The Advanced Message Queuing Protocol (AMQP) is an open standard application layer protocol for message-oriented middleware. The defining features of AMQP are message orientation, queuing, routing (including point-to-point and publish-and-subscribe), reliability and security.

AMQP mandates the behaviour of the messaging provider and client to the extent that implementations from different vendors are truly interoperable, in the same way as SMTP, HTTP, FTP, etc., have created interoperable systems. Previous attempts to standardize middleware have happened at the API level (e.g. JMS) and this did not create interoperability[2]. Unlike JMS, which merely defines an API, AMQP is a wire-level protocol. A wire-level protocol is a description of the format of the data that is sent across the network as a stream of octets. Consequently any tool that can create and interpret messages that conform to this data format can interoperate with any other compliant tool irrespective of implementation language.

5.4.1.2 JMS

The Java Message Service (JMS) API is a Java Message Oriented Middleware (MOM) API for sending messages between two or more clients. JMS is a part of the Java Platform, Enterprise Edition, and is defined by a specification developed under the Java Community Process as JSR 914. It is a messaging standard that allows application components based on the Java Enterprise Edition (JEE) to create, send, receive, and read messages. It allows the communication between different components of a distributed application to be loosely coupled, reliable, and asynchronous.

5.4.1.3 STOMP

Simple (or Streaming) Text Oriented Message Protocol (STOMP), formerly known as TTMP, is a simple text-based protocol, designed for working with Message Oriented Middleware. It provides an interoperable wire format that allows STOMP clients to talk with any Message Broker supporting the protocol. It is thus language-agnostic, meaning a broker developed for one language or platform can receive communications from client software developed in another language.

5.4.2 Rabbit MQ

RabbitMQ (http://www.rabbitmq.com/) is an open source message broker software (message-oriented middleware), using the standard Advanced Message Queuing Protocol (AMQP). The RabbitMQ server is written in Erlang and is built on the Open Telecom Platform framework for clustering and failover. The RabbitMQ project consists of several parts:

- The RabbitMQ exchange server itself
- Gateways for HTTP, XMPP and STOMP protocols
- AMQP client libraries for Java and .NET Framework. (AMQP clients for other languages are available from other vendors)
- The "Shovel" plug-in that takes care of copying (replicating) messages from one broker to another.

Its principal idea is pretty simple: it accepts and forwards messages. In the following Figure 5-1 depicts the simplest queue that can be administrated by RabbitMQ. A program that sends a message is a producer, denoted with the letter P, while a program that receives messages is a consumer, denoted with the letter C. As we can see in Figure 5-1 a producer produces a message that is stored to the queue in order to be processed by the consumer whenever it is available.

Figure 5-1 Simplest queue

In our scenario a slightly more complex queue is used that aims at improving scalability. In this case (Figure 5-2) a work queue is created that is used to distribute time-consuming tasks among multiple workers. More precisely the producer queues XML documents that are going to be RDFized, enriched and stored in the queue. In the following the work queue distributes these documents to a number of consumers that do the data manipulation concurrently and store the results.

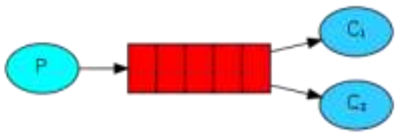

Figure 5-2 Work queue

Moreover a remote procedure call (RPC) is used. RPC (Figure 5-3) is an inter-process communication that allows a computer program to cause a subroutine or procedure to execute in another address space (commonly on another computer on a shared network) without the programmer explicitly coding the details for this remote interaction. That is, the programmer writes essentially the same code whether the subroutine is local to the executing program, or remote. The main idea is that the client sends a request message to a known remote server to execute a specified procedure with supplied parameters. In the following the remote server sends a response to the client, and the application continues its process.

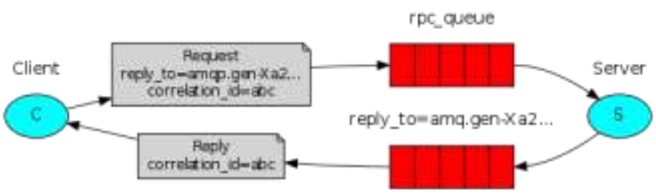

Figure 5-3 Remote Procedure Call

5.4.3 MINT PI

MINT PI is part of the Mint interoperability services suite (*http://mint.image.ece.ntua.gr*) and serves mainly as a scalable mechanism for structured data processing. It is built using RabbitMQ in its core and utilizes a number of message queue patterns as they were described earlier. The approach of using message queues instead of a mechanism like Hadoop was decided based on the requirement to scale on both large and small datasets without reducing the efficiency of the overall system. The overall architecture of MINT PI is depicted in *Figure 4*-4.

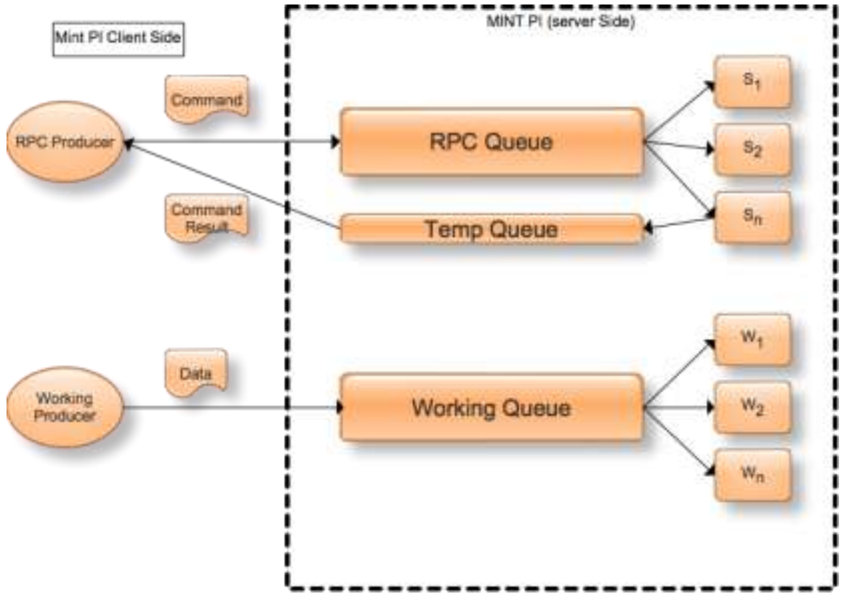

Figure 4-4 The Mint PI Architecture

MINT PI is using two distinct queue patterns, an RPC Queue pattern which is used for cases where the client desires to block while the processing is executed and also awaits for a response in a pre-defined format, and a Working Queue pattern which is used for non blocking processing where the client submits the data for processing and does not wait for a response. The first case is mainly used for the implementation of specific commands, e.g. for implementing the cleaning or deletion of a repository, while the second case is used for bulk processing of raw data, e.g. data transformation and enrichment of records.

Scalability is achieved by the parallel processing of many workers for the case of the working Queue and the existence of number of RPC Consumers that are running concurrently for the case of the RPC Queue. In both cases the workers and the RPC consumers might be running on different nodes of the cluster that is materialized while more workers and consumers can be added to the system at any time and thus increasing the processing power of the system. At the same time, the RabbitMQ broker is also scalable, new nodes can be added to the broker and thus increasing the message per second ratio that can be handled by the system. In this way an overall scalable architecture for message delivery is defined which can be extended with minimal administrative effort.

MINT PI is not limited to a certain type of processing or data schema that is delivered using the messages. This is achieved by utilizing a software design pattern named the Strategy or Policy Pattern, using this pattern it is possible to select different algorithms for execution on runtime. A UML diagram that represents the Strategy or Policy pattern is depicted in *Figure 5*. Another benefit from using this design pattern is the abstraction that is introduced between the messaging layer of the system and the implementation of various algorithms for processing. A developer does not have to know about the intrinsic details of the messaging system in order to implement another algorithm for data processing by Mint PI.

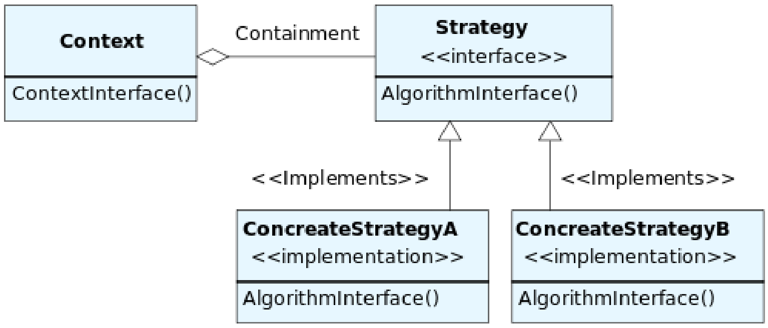

Figure 5 The Strategy or Policy Design Pattern

By following the architectural principles of Mint PI, a number of strategies have been implemented in order to support the semantic enrichment of data. More specifically the following strategies have been developed in order to support the functionalities defined by the Semantic Search Pilot of the Indicate Project:

1. A strategy for data transformation(this one was also used to perform the RDFization of data)
2. A strategy for data enrichment, and finally
3. A strategy for storing the final data into a semantic repository.

In the same sense a number of commands have been defined again as strategies but for the RPC Queue, for performing a number of commands to the semantic repository, e.g. deletion of records.

The deployment of Mint PI on Amazon EC2 was a straight forward process. A number of server images was created to support the different parts that combined define the Mint PI architecture. These Server images are the following:

1. The first image contains the RabbitMQ broker which can also be used to add more nodes to it if it is required by the load of the cluster.
2. Another image was created containing an initial predefined number of both Workers and RPC consumers. The number of workers and consumers can be adjusted on runtime on the same running image on the cloud or, if it is needed, a new node can be initialized and thus increase the number of available workers and consumers on the system.

The overall deployed architecture is characterized by elasticity in the sense that the resources available, either as broker nodes or worker/consumer nodes, can be adjusted at any time depending on the load that the system has to handle.

6 Semantic Search Evaluation

Our main objective was to establish a semantic search system on e-Infrastructures using the MICHAEL metadata and to present the main benefits from this architecture. Towards this aim, MICHAEL metadata were enhanced by the semantic transformation and the linking with other sources. Furthermore SPARQL support provided by the semantic repository and in that way expressive semantic retrieval is supported. The simple case scenario that is to use only information explicitly defined in the MICHAEL dataset is retained so one could search for items made in Italy. The main advantage though of this architecture is the support of more complex semantic queries that permit one to search for all the items that are hold by countries of the Mediterranean. In this query the concept Mediterranean Sea, that describes all the country resources of the Mediterranean Sea, is used in a SPARQL query because of the linking of MICHAEL countries to DBPedia. In a similar manner, expressive concepts having a semantic interpretation like politicians; living persons and other can be combined in a query for the retrieval of MICHAEL content.

In order to illustrate the improved performance of semantic search on e-infrastructures we compared it with two other architectures. The first one, called Single Host, consists of a single thread process on local host that is used for the RDFization and enrichment of the data as well as a 4store triplestore also on a local host used for the storage of the data. The second architecture, called Local Cluster, uses a multi-thread process on a local cluster for the RDFization and enrichment of the data and a 4store triplestore also on a local cluster for the storage of the data. Finally, Amazon Cloud is the architecture that uses a multi-thread process on the Amazon cloud in order to perform the RDFization and enrichment and a 4store triplestore, also on the Amazon cloud, for storing the data. The data that have been used are MICHAEL data.

Table 3 illustrates the total time required by each architecture in order to perform the RDFization, enrichment and storage of the MICHAEL dataset that contain 8511 items. We can see that the Local Host needed approximately 6,2 hours to perform the task, the Local Cluster about 1,39 hours while the Amazon Cloud needed only 23,7 minutes. We observe that Amazon cloud outperforms the other two architectures by being faster for several orders of magnitude. So we can conclude that e-Infrastructures improve to a great extend the efficiency of our system. Finally it is worth noting that this vast difference of the proposed three architectures can be increased even more when considering a larger data set such as Europeana that contains approximately 20 million items.

Method Used	Time in Millisecs
Local Host	22.383.937ms (~6.2hrs)
Local Cluster	5.020.430 (~1.39hrs)
Amazon Cloud	1.422.000 (~23.7 min)

Table 3 Architecture evaluation

7 Conclusions

This deliverable describes the pilot of the semantic search using e-infrastructures. For the implementation of this pilot the MICHAEL dataset has been used. The first step was to transform the metadata to RDF that can provide a semantically richer representation. After that the values of specific elements of the dataset like countries, persons and languages were externally linked using DBpedia. The transformed to RDF and enriched metadata that resulted from the previous modules were then stored to semantic repository 4store that fulfilled the requirements of SPARQL query language, distributed deployment and free licence. The very important outcome of this procedure was the support of queries for retrieval of MICHAEL content that were using information that was not available on the original dataset.

The aforementioned workflow was deployed on Amazon Elastic Compute Cloud (EC2) that is considered one of the most mature Cloud e-infrastructures at the moment. For the deployment on Amazon Elastic Compute Cloud a processing interface (MINT-PI) was implemented that parallelizes the semantic transformation and enrichment taking in that way advantage of the massive processing power that can be offered by a Cloud infrastructure.

Evaluation was performed for the enrichment process as well as for the overall performance of the proposed architecture. More specifically, the algorithm that was developed for the resource discovery operated outstandingly well for the countries and the languages, while it also estabished very good performance for persons, considering the difficulty of the task. More specifically the main diffuculty with the linking of person resources is that there is no guarantee that the resource can be discovered in the external source. In other words, the fact that a person is found in the dataset does not necessarily make it a resource served by external sources, while on the other hand countries and languages resources exist in more than one external sources making their discovery easier. Finally, comparing the overall performance of the workflow when deployed on cloud e-infrastructures and when deployed on a single server it was observed that e- infrastructures can provide scalability that is an essential feature when processing massive volume of content –that is always the case for digital cultural heritage - with very low cost. In particular the cost for the processing of the MICAHEL data cost $ 0.68 per node per hour while the storaging of 1 Gb per month costs $0.11. In other words the overall cost for the deployment of the INDICATE semantic search pilot was ~ 6.1 € (~4.4€ for storing + 1.7€ for processing) that is extremely low for the quality of the offered services.

8 References

[1] http://www.cordis.europa.eu/pub/ist/docs/digicult/lund

[2] SIEDL: First Workshop on Semantic Interoperability in the European Digital Library, 5th European Semantic Web Conference, Tenerife, Spain, June 2, 2008

[3] http://www.europeana.eu

[4] Klyne, G. & Carroll, J. (2004). "Resource Description Framework (RDF): Concepts and Abstract Syntax" - W3C Recommendation. Retrieved June 14, 2009, Consulted January 31, 2012. Available: http://www.w3.org/TR/rdfconcepts/

[5] McGuinness, D. L. & Harmelen, F.V (2004) "OWL Web Ontology Language Overview - W3C Recommendation", Consulted January 31, 2012. Available: http://www.w3.org/TR/2004/REC-owl-features-20040210/

[6] http://www.MICHAEL-culture.org/en/about/project

[7] Prud'hommeaux, E. & Seaborne, A. (2008) "SPARQL Query Language for RDF - W3C Recommendation", Consulted January 31, 2012. Available: http://www.w3.org/TR/rdf-sparql-query/

[8] Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm

[9] http://dbpedia.org/About

[10] http://4store.org/

[11] http://www.openrdf.org/

[12] http://www.ontotext.com/OWLIM

[13] Bizer, C. & Heath, T. & Berners-Lee, T. (2009) "Linked Data - The Story So Far", International Journal on Semantic Web and Information Systems, Volume: 5, Issue: 3, Publisher: Elsevier, Pages: 1-22 ISSN: 15526283, DOI: 10.4018/jswis.2009081901

[14] Berners-Lee, T. (2006). "Linked Data - Design Issues", Consulted January 31, 2012. Available: http://www.w3.org/DesignIssues/LinkedData.htm

[15] http://www.w3.org/2009/08/skos-reference/skos.html

[16] Brickley, D. & Guha, R. V.(2004) "RDF Vocabulary Description Language 1.0: RDF Schema - W3C Recommendation", Consulted January 31, 2012. Available: http://www.w3.org/TR/rdf-schema/

[17] Dan Brickley, Libby Miller, 2010, http://www.foaf-project.org/

[18] Ian Davis, Eric Vitiello Jr, 2004, http://purl.org/vocab/relationship/

[19] DCMI Usage Board, 2010, http://purl.org/dc/dcmitype/

[20] Tim Cole, Tom Habing, Jane Hunter, Pete Johnston, Carl Lagoze, 2003, http://purl.org/dc/elements/1.1/

[21] Tim Cole, Tom Habing, Jane Hunter, Pete Johnston, Carl Lagoze, 2006, http://purl.org/dc/terms/

[22] Edd Dumbill, 2002, http://lod.taxonconcept.org/ontology/txn.owl

[23] Samer A. Abdallah, Bob Ferris, 2010, http://purl.org/ontology/olo/core

[24] Yves Raimond, Mats Skillingstad, Bob Ferris, 2010, http://purl.org/ontology/co/core

[25] Heath, T. & Bizer, C (2011) "Linked Data: Evolving the Web into a Global Data Space" (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool DOI: 10.2200/S00334ED1V01Y201102WBE001 ISBN: 9781608454303 (paperback) ISBN: 9781608454310 (ebook)

[26] http://www.systap.com/bigdata.html

[27] http://www.dist-systems.bbn.com/people/krohloff/shard.shtml

[28] http://docs.dydra.com/dydra

9 Appendix

9.1 XML File


```
<?xml version="1.0" encoding="UTF-8"?><digital-collection id="UK-DC-7036a5f"><identification><title>London
Metropolitan Archives: schooLMAte</title></identification><description><legal-status>Rights Reserved</legal-
status><language-group><language code="eng"
scheme="michael:/portal/languages"><cs>Angličtina</cs><ee>Inglise</ee><en>English</en><pl>English</pl><fr>anglais<
/fr><it>English</it><fi>englanti</fi><sv>engelska</sv><de>Englisch</de><el>Αγγλικά</el><nl>Engels</nl><hu>Angol</h
u><bg>английски</bg><sk>anglicky</sk><es>inglés</es></language><language code="fra"
scheme="michael:/portal/languages"><cs>Francouzština</cs><ee>Prantsuse</ee><en>French</en><pl>French</pl><fr>franç
ais</fr><it>French</it><fi>ranska</fi><sv>franska</sv><de>Französisch</de><el>Γαλλικά</el><nl>Frans</nl><hu>Franci
a</hu><bg>френски</bg><sk>francúzsky</sk><es>francés</es></language><language
xml:lang="en">Irish</language><language xml:lang="cy">Gwyddelig</language></language-group><digital-format-
group><digital-format code="html" scheme="michael:/portal/digital-
formats"><fr>HTML</fr><el>HTML</el><es>HTML</es><pl>HTML</pl><de>HTML</de><hu>HTML</hu><lv>HTML</lv><nl>HTML</nl><
fi>HTML</fi><en>HTML</en><it>HTML</it><sv>HTML</sv><cy>HTML</cy><cs>HTML</cs><ee>HTML</ee><sk>html</sk></digital-
format><digital-format code="jpeg" scheme="michael:/portal/digital-
formats"><fr>JPEG</fr><el>JPEG</el><es>JPEG</es><pl>JPEG</pl><hu>JPEG</hu><lv>JPEG</lv><nl>JPEG</nl><fi>JPEG</fi><
en>JPEG</en><cy>JPEG</cy><sv>JPEG</sv><it>JPEG</it><cs>JPEG</cs><ee>JPEG</ee><de>JPEG</de><sk>jpeg</sk></digital-
format><digital-format code="flash" scheme="michael:/portal/digital-formats"><fr>Macromédia
Flash</fr><en>Macromedia Flash</en><it>Macromedia Flash</it><fi>Macromedia Flash</fi><es>Macromedia
Flash</es><pl>Macromedia Flash</pl><de>Macromedia Flash</de><hu>Macromedia Flash</hu><nl>Macromedia
Flash</nl><cy>Macromedia Flash</cy><sv>Macromedia Flash</sv><cs>Macromedia Flash</cs><ee>Macromedia
Flash</ee><el>Flash</el><lv>Macromédia Flash</lv><sk>makromédia flash</sk></digital-format><digital-format
code="real" scheme="michael:/portal/digital-formats"><en>Real Audio</en><cy>Real Audio</cy><fi>Real
Audio</fi><es>Real Audio</es><pl>Real Audio</pl><hu>Real Audio</hu><lv>Real Audio</lv><nl>Real Audio</nl><el>Real
Audio</el><it>Real Audio</it><sv>Real Audio</sv><fr>Real Audio</fr><cs>Real Audio</cs><ee>Real Audio</ee><de>Real
Audio</de><sk>real audio</sk></digital-format></digital-format-group><digital-type-group><digital-type
code="dt001.1" scheme="michael:/portal/digital-
types"><fr>Texte</fr><en>Text</en><cy>Testun</cy><it>Testo</it><cs>Text</cs><sk>Text</sk><sv>Text</sv><ee>Tekst</e
e><pl>Tekst</pl><nl>Tekst</nl><fi>Tekstiaineistoa</fi><lv>Teksts</lv><hu>Szöveg</hu><el>Κείμενο</el><de>Text</de><
es>Texto</es></digital-type><digital-type code="dt001.2" scheme="michael:/portal/digital-types"><fr>Image
fixe</fr><en>Still image</en><cy>Delwedd lonydd</cy><it>Immagine
fissa</it><cs>Obrázek</cs><ee>Kujutis</ee><fi>Stillkuvia</fi><sv>Stillbilder</sv><nl>Statische
afbeelding</nl><lv>Attēls</lv><hu>Állókép</hu><el>Σταθερή εικόνα</el><de>Bild</de><pl>Obraz
nieruchomy</pl><sk>Obrázok</sk><es>Imagen fija</es></digital-type><digital-type code="dt001.2.2"
scheme="michael:/portal/digital-types"><fr>Cartographie (raster)</fr><en>Map (raster image)</en><cy>Map (delwedd
raster)</cy><it>Cartografia (raster)</it><cs>Bitmapa (rastrový obrázek)</cs><ee>Rasterkujutis</ee><fi>Kartta
(rasteri kuva)</fi><sv>Karta (raster)</sv><nl>Cartografie (raster)</nl><lv>Karte</lv><hu>Térkép</hu><de>Karte
(Rasterbild)</de><el>Χάρτης (raster)</el><pl>Mapa (obraz rastrowy)</pl><sk>Bitmapa (rastrový obrázok)</sk><es>Mapa
(imagen raster)</es></digital-type><digital-type code="dt001.4" scheme="michael:/portal/digital-
types"><fr>Son</fr><en>Sound
files</en><en>Sound</en><cy>Swn</cy><it>Suono</it><cs>Zvuk</cs><sk>Zvuk</sk><ee>Heli</ee><fi>Äänitiedostoja</fi><s
v>Ljud</sv><nl>Geluid</nl><lv>Skaņa</lv><hu>Hang</hu><de>Ton</de><el>Αρχείο
ήχου</el><pl>Dźwięk</pl><es>Sonido</es></digital-type><digital-type code="dt001.7"
scheme="michael:/portal/digital-types"><fr>Ressource interactive</fr><en>Interactive resource</en><cy>Adnodd
rhyngweithiol</cy><it>Risorsa interattiva</it><cs>Interaktivní zdroj</cs><ee>Interaktiivne
vahend</ee><fi>Interaktiivinen resurssi</fi><sv>Interaktiv resurs</sv><nl>Interactieve bron</nl><lv>Interaktīvs
resurss</lv><hu>Interaktív forrás</hu><el>Διαδραστικός πόρος</el><de>Interaktive Ressource</de><pl>Zasoby
interaktywne</pl><sk>Interaktívny zdroj</sk><es>Recurso interactivo</es></digital-type></digital-type-
group><size>300 images</size><accrual>complete</accrual><access-control>unrestricted</access-
control><abstract><p>This collection of digital images, interpretative text, maps and audioclips allows children
and teachers to explore the complex histories of: Black and Asian Londoners; the French community; the Irish
community; and the communities of the City of London.  The resources draw on important photographs, designs,
letters, prints, maps and interviews to illustrate the important contribution that these communities have made to
the making of London from 1067 to 2003.</p></abstract></description><subject-indexing><subject-group><subject
code="76" scheme="michael:/portal/subjects"><cs>Kulturní identita</cs><ee>Kultuuriline identiteet</ee><en>Cultural
identity</en><it>Identità culturale</it><fr>Identité culturelle</fr><cy>Hunaniaeth
ddiwylliannol</cy><fi>Kulttuuri-identiteetti</fi><sv>Kulturell identitet</sv><el>Πολιτιστική
ταυτότητα</el><nl>Culturele identiteit</nl><lv>Kultūras identitāte</lv><hu>Kulturális identitás</hu><de>Kulturelle
Identität</de><pl>Tożsamość kulturowa</pl><sk>Kultúrna identita</sk><es>Identidad cultural</es></subject><subject
code="88"
scheme="michael:/portal/subjects"><cs>Náboženství</cs><ee>Religioon</ee><en>Religion</en><it>Religione</it><fr>Rel
igion</fr><sv>Religion</sv><cy>Crefydd</cy><fi>Uskonto</fi><el>Θρησκεία</el><nl>Religie</nl><lv>Reliģija</lv><hu>V
allás</hu><de>Religion</de><bg>Религия</bg><pl>Religia</pl><sk>Náboženstvo</sk><es>Religión</es></subject><subject
code="96"
scheme="michael:/portal/subjects"><cs>Historie</cs><ee>Ajalugu</ee><en>History</en><it>Storia</it><fr>Histoire</fr
><cy>Hanes</cy><fi>Historia</fi><es>Historia</es><pl>Historia</pl><sv>Historia</sv><el>Ιστορία</el><nl>Geschiedeni
s</nl><lv>Vēsture</lv><hu>Történelem</hu><de>Geschichte</de><bg>история</bg><sk>História</sk></subject><subject
code="290" scheme="michael:/portal/subjects"><cs>Pracovní prostředí</cs><ee>Töökeskkond</ee><en>Work
environment</en><it>Ambiente di lavoro</it><fr>Milieu du
travail</fr><fi>Työympäristö</fi><sv>Arbetsmiljö</sv><el>Περιβάλλον εργασίας</el><nl>Arbeidsmilieu</nl><lv>Darba
vide</lv><hu>Munkakörnyezet</hu><de>Arbeitsleben</de><bg>Работна среда</bg><pl>Środowisko pracy</pl><sk>Pracovné
prostredie</sk><es>Ambiente de trabajo</es></subject><subject code="282"
scheme="michael:/portal/subjects"><cs>Doprava</cs><ee>Transport</ee><pl>Transport</pl><sv>Transport</sv><en>Transp
ort</en><it>Trasporti</it><fr>Transport</fr><cy>Trafnidiaeth</cy><fi>Kuljetusala</fi><el>Συγκοινωνία</el><nl>Vervo
```

er</nl><lv>Transports</lv><hu>Közlekedés</hu><de>Transportgeschichte</de><bg>Транспорт</bg><sk>doprava</sk><es>Tra
nsporte</es></subject><subject code="211"
scheme="michael:/portal/subjects"><cs>Migrace</cs><ee>Migratsioon</ee><en>Migration</en><it>Migrazione</it><fr>Mig
ration</fr><sv>Migration</sv><fi>Siirtolaisuus</fi><el>Μετανάστευση</el><cy>Hanes
ymfudo</cy><nl>Migratie</nl><lv>Migrācija</lv><hu>Migráció</hu><de>Migrationsgeschichte</de><bg>Миграция</bg><pl>M
igracja</pl><sk>Migrácia</sk><es>Migración</es></subject><subject code="244"
scheme="michael:/portal/subjects"><cs>Dějiny práva</cs><ee>Õiguse ajalugu</ee><en>Legal history</en><it>Storia del
diritto</it><fr>Histoire du droit</fr><fi>Oikeushistoria</fi><sv>Rättshistoria</sv><el>Ιστορία του
δικαίου</el><nl>Rechtsgeschiedenis</nl><lv>Tiesību
vēsture</lv><hu>Jogtörténet</hu><de>Rechtsgeschichte</de><bg>история на правото</bg><pl>Historia
prawa</pl><sk>Dejiny práva</sk><es>Historia del derecho</es></subject><subject code="150"
scheme="michael:/portal/subjects"><cs>Grafika</cs><sk>Grafika</sk><hu>Grafika</hu><ee>Graafika</ee><en>Graphic
arts</en><fr>Arts graphiques</fr><it>Grafica</it><cy>Celfyddydau
graffig</cy><fi>Grafiikka</fi><sv>Grafik</sv><el>Γραφιστικές τέχνες</el><nl>Grafische kunsten</nl><lv>Grafiskā
māksla</lv><de>Grafik</de><bg>изобразително изкуство</bg><pl>Grafika artystyczna</pl><es>Artes
gráficas</es></subject><subject code="179" scheme="michael:/portal/subjects"><cs>Volný čas</cs><ee>Vaba
aeg</ee><en>Leisure</en><it>tempo libero</it><fr>Loisir</fr><cy>Hamdden</cy><fi>Urheilu ja vapaa-
aika</fi><sv>Fritid</sv><el>Ελεύθερος χρόνος</el><nl>Vrijetijd</nl><lv>Brīvais
laiks</lv><hu>Szórakozás</hu><de>Freizeit</de><bg>Свободно време</bg><pl>Rekreacja</pl><sk>Voľný
čas</sk><es>Ocio</es></subject><subject code="204" scheme="michael:/portal/subjects"><cs>Sociální politika a
sociální péče</cs><ee>Sotsiaalpoliitika ja hoolekanne</ee><en>Social policy and welfare</en><it>Politica sociale e
benessere</it><fr>Politique et bien-être social</fr><cy>Polisi a lles
cymdeithasol</cy><fi>Sosiaalipolitiikka</fi><sv>Socialpolitik och omsorg</sv><el>Κοινωνική πολιτική και
πρόνοια</el><nl>Sociale politiek en welzijn</nl><lv>Sociālā politika un labklājība</lv><hu>Szociális-jóléti
stratégia</hu><de>Sozialpolitik und Wohlfahrt</de><bg>Социална политика</bg><pl>Polityka społeczna i
zdrowotna</pl><sk>Sociálna politika a sociálna starostlivosť</sk><es>Política y bienestar
social</es></subject><subject code="212" scheme="michael:/portal/subjects"><cs>Etnické menšiny</cs><sk>Etnické
menšiny</sk><ee>Mustad ja etnilised vähemused</ee><en>Black and ethnic minority groups</en><it>Minoranze
etniche</it><fr>Minorité ethnique</fr><cy>Grwpiau pobl dduon a lleiafrifoedd
ethnig</cy><fi>Vähemmistöt</fi><sv>Färgade och etniska minoriteter</sv><el>Μειονοτικές και εθνοτικές
ομάδες</el><nl>Etnische minderheden</nl><lv>Etnisko minoritāšu grupas</lv><hu>Etnikumok,
kisebbségek</hu><de>Ethnische Minderheiten</de><bg>Малцинства</bg><pl>Mniejszości etniczne</pl><es>Grupos
minoritarios étnicos</es></subject><subject code="225" scheme="michael:/portal/subjects"><cs>Rozvoj
měst</cs><ee>Linna-areng</ee><en>Urban development</en><it>Sviluppo urbano</it><fr>Développement
urbain</fr><cy>Datblygu trefol</cy><fi>Kaupunkisuunnittelu</fi><sv>Urban utveckling</sv><el>Αστική
ανάπτυξη</el><nl>Stedelijke ontwikkeling</nl><lv>Pilsētu
attīstība</lv><hu>Városfejlődés</hu><de>Stadtentwicklung</de><bg>Урбанистично развитие</bg><pl>Rozwój
miast</pl><sk>Rozvoj miest</sk><es>Desarrollo urbano</es></subject><subject code="215"
scheme="michael:/portal/subjects"><cs>Stavby a urbanismus</cs><ee>Ehitised ja linnakeskkond</ee><en>Buildings and
the built environment</en><en>Buildings and the build environment</en><it>Edifici e paesaggio
urbano</it><fr>Bâtiment et urbanisme</fr><fi>Rakennukset ja asuinympäristö</fi><sv>Byggnader och
byggnadsmiljö</sv><el>Κτίρια και δομημένο περιβάλλον</el><cy>Adeiladau a'r amgylchedd adeiledig</cy><nl>Gebouwen
en de gebouwde omgeving</nl><lv>Celtnes un to vide</lv><hu>Épületek, épített környezet</hu><de>Bauwerke,
Landschaftsarchitektur, Stadt-, Verkehrs- und Raumplanung</de><bg>Сгради и околна среда</bg><pl>Budowle i
środowisko budowlane</pl><sk>Stavby a urbanizmus</sk><es>Edificios y el ambiente construido</es></subject><subject
code="64" scheme="michael:/portal/subjects"><cs>Zdraví a životní styl</cs><ee>Elustiil</ee><en>Health &
lifestyle</en><it>Salute e stile di vita</it><fr>Santé et style de vie</fr><cy>Iechyd a ffordd o
fyw</cy><fi>Terveys ja elintavat</fi><sv>Hälsa och livsstil</sv><el>Υγεία και τρόπος ζωής</el><nl>Gezondheid &
levensstijl</nl><lv>Veselība un dzīvesstils</lv><hu>Egészség és életmód</hu><de>Gesundheit &
Lebensstil</de><pl>Zdrowie i styl życia</pl><sk>Zdravie a životný štýl</sk><es>Salud y estilo de
vida</es></subject><subject code="259" scheme="michael:/portal/subjects"><cs>Politika a vláda</cs><sk>Politika a
vláda</sk><ee>Poliitika ja valitsus</ee><en>Politics and government</en><it>Politica</it><fr>Politique et
gouvernement</fr><cy>Gwleidyddiaeth a'r llywodraeth</cy><fi>Politiikka ja valtio</fi><sv>Politik och
myndigheter</sv><el>Πολιτική και κυβέρνηση</el><nl>Politiek en bestuur</nl><lv>Politika un
pārvalde</lv><hu>Politika és kormány</hu><de>Politik und Regierung</de><bg>политика и
правителство</bg><pl>Polityka i rząd</pl><es>Política y gobierno</es></subject><subject code="275"
scheme="michael:/portal/subjects"><cs>Stavební, vojenské a důlní strojírenství</cs><ee>Tsiviil-, sõjandus- ja
kaevandamistehnika</ee><en>Civil, military and mining engineering</en><it>Ingegneria civile, militare e delle
miniere</it><it>Ingegneria civile, militare e mineraria</it><fr>Génie civil, militaire et
minier</fr><cy>Peirianneg sifil, milwrol a mwyngloddio</cy><fi>Kone- ja rakennusteollisuus</fi><sv>Civil och
militär ingenjörsvetenskap</sv><el>Πολιτική, στρατιωτική και μεταλλευτική μηχανική</el><nl>Civiele, militaire en
mijnbouwtechniek</nl><lv>Civilā un militārā inženierija, kalnrūpniecība</lv><hu>Civil, katonai és bányászati
mérnökség</hu><de>Ingenieurwesen, Militärtechnik und Bergbau</de><bg>Цивилно, военно и минно
инженерство</bg><pl>Inżynieria cywilna, wojskowa i wydobywcza</pl><sk>Stavebné, vojenské a banícke
inžinierstvo</sk><es>Ingeniería civil, militar y de minas</es></subject></subject-group><period-group><period
code="med" scheme="michael:/portal/periods"><ee>Keskaeg</ee><cs>Středověk</cs><fr>Moyen
Âge</fr><en>Medieval</en><it>Medioevo</it><fi>Keskiaika</fi><sv>Medeltid</sv><el>Μεσαιωνική
περίοδος</el><nl>Middeleeuwen</nl><lv>Viduslaiki</lv><hu>Középkor</hu><de>Mittelalter</de><bg>Средновековие</bg><c
y>Canoloesol</cy><pl>Średniowiecze</pl><sk>stredovek</sk><es>Edad Media</es></period><period code="latemed"
scheme="michael:/portal/periods"><ee>Kõrg- ja hiliskeskaeg</ee><cs>Vrcholný a pozdní středověk</cs><fr>Age Féodal
et Bas Moyen Âge</fr><en>High and late medieval</en><it>Basso medioevo</it><fi>Myöhäiskeskiaika</fi><sv>Hög- och
sen-medeltid</sv><el>Μέση και Ύστερη Μεσαιωνική περίοδος</el><nl>Hoge en late Middeleeuwen</nl><lv>Attīstītie un
vēlie viduslaiki</lv><hu>Középkor (11-15. sz.)</hu><de>Hoch- und Spätmittelalter</de><bg>Средно и късно
Средновековие</bg><pl>Środkowe i późne Średniowiecze</pl><sk>Vrcholný a neskorý stredovek</sk><es>Baja Edad
Media</es></period><period code="c16" scheme="michael:/portal/periods"><ee>16. sajand</ee><cs>16.
století</cs><fr>16e siècle</fr><en>16th century</en><it>XVI secolo</it><fi>1500 -luku</fi><sv>1500-
talet</sv><el>16ος αιώνας</el><nl>16e eeuw</nl><lv>16. gadsimts</lv><hu>16. század</hu><de>16.
Jahrhundert</de><bg>16 век</bg><cy>16eg ganrif</cy><pl>XVI wiek</pl><sk>16.storočie</sk><es>Siglo
XVI</es></period><period code="c17" scheme="michael:/portal/periods"><ee>17. sajand</ee><cs>17.
století</cs><fr>17e siècle</fr><en>17th century</en><it>XVII secolo</it><fi>1600 -luku</fi><sv>1600-
talet</sv><el>17ος αιώνας</el><nl>17e eeuw</nl><lv>17. gadsimts</lv><hu>17. század</hu><de>17.
Jahrhundert</de><bg>17 век</bg><cy>17eg ganrif</cy><pl>XVII wiek</pl><sk>17.storočie</sk><es>Siglo
XVII</es></period><period code="c18" scheme="michael:/portal/periods"><ee>18. sajand</ee><cs>18.
století</cs><fr>18e siècle</fr><en>18th century</en><it>XVIII secolo</it><fi>1700 -luku</fi><sv>1700-
talet</sv><el>18ος αιώνας</el><nl>18e eeuw</nl><lv>18. gadsimts</lv><hu>18. század</hu><de>18.

Jahrhundert</de><bg>18 век</bg><cy>18fed ganrif</cy><pl>XVIII wiek</pl><sk>18.storočie</sk><es>Siglo XVIII</es></period><period code="c19" scheme="michael:/portal/periods"><ee>19. sajand</ee><cs>19. století</cs><fr>19e siècle</fr><en>19th century</en><it>XIX secolo</it><fi>1800 -luku</fi><sv>1800-talet</sv><el>19ος αιώνας</el><nl>19e eeuw</nl><lv>19. gadsimts</lv><hu>19. század</hu><de>19. Jahrhundert</de><bg>19 век</bg><cy>19eg ganrif</cy><pl>XIX wiek</pl><sk>19.storočie</sk><es>Siglo XIX</es></period><period code="c20" scheme="michael:/portal/periods"><ee>20. sajand</ee><cs>20. století</cs><fr>20e siècle</fr><en>20th century</en><it>XX secolo</it><fi>1900 -luku</fi><sv>1900-talet</sv><el>20ος αιώνας</el><nl>20e eeuw</nl><lv>20. gadsimts</lv><hu>20. század</hu><de>20. Jahrhundert</de><bg>20 век</bg><cy>20fed ganrif</cy><pl>XX wiek</pl><sk>20.storočie</sk><es>Siglo XX</es></period><period code="c21" scheme="michael:/portal/periods"><ee>21. sajand</ee><cs>21. století</cs><fr>21e siècle</fr><en>21st century</en><it>XXI secolo</it><fi>2000 -luku</fi><sv>2000-talet</sv><el>Σύγχρονη εποχή (21ος αι. μ.Χ).</el><nl>21e eeuw</nl><lv>21. gadsimts</lv><hu>21. század</hu><de>21. Jahrhundert</de><bg>21 век</bg><cy>21ain ganrif</cy><pl>XXI wiek</pl><sk>21.storočie</sk><es>Siglo XXI</es></period></period-group><culture-group><culture code="weu" scheme="michael:/portal/civilisations"><cs>Západoevropská</cs><en>Western European</en><es>Europa Occidental</es><cy>Gorllewin Ewrop</cy><fi>Länsi-Eurooppalainen</fi><sv>Västeuropeisk</sv><nl>West-Europese</nl><hu>Nyugat-európai</hu><pl>Zachodnioeuropejska</pl><sv>Västeuropeisk</sv></culture></culture-group><spatial-coverage-group><spatial-coverage code="gbe" scheme="michael:/portal/spatial-coverages"><cs>Anglie</cs><ee>Inglismaa</ee><en>England</en><it>Inghilterra</it><fr>Angleterre</fr><fi>Englanti</fi><sv>England</sv><nl>Engeland</nl><lv>Anglija</lv><cy>Lloegr</cy><sk>Anglicko</sk><es>Inglaterra</es></spatial-coverage><spatial-coverage code="gb" scheme="michael:/portal/spatial-coverages"><cs>SPOJENÉ KRÁLOVSTVÍ (VELKÁ BRITÁNIE)</cs><ee>Ühendkuningriik</ee><en>UNITED KINGDOM</en><cy>Y DEYRNAS UNEDIG</cy><it>REGNO UNITO</it><fr>ROYAUME UNI</fr><fi>Iso-Britannia</fi><sv>Storbrittanien</sv><el>ΗΝΩΜΕΝΟ ΒΑΣΙΛΕΙΟ</el><nl>VERENIGD KONINKRIJK</nl><lv>APVIENOTĀ KARALISTE</lv><de>VEREINIGTES KÖNIGREICH (UK)</de><sk>Spojené kráľovstvo</sk><es>REINO UNIDO</es></spatial-coverage></spatial-coverage-group><start-date>1067</start-date><end-date>2003</end-date><famous-event-group><famous-event>Great Fire of London, 1666</famous-event><famous-event>Great Plague of London, 1665</famous-event><famous-event>Black Death, 1348</famous-event></famous-event-group><famous-place-group><famous-place>Guildhall, London</famous-place><famous-place>Westminster Bridge, London</famous-place><famous-place>Old London Bridge</famous-place><famous-place>Monument, London</famous-place><famous-place>River Thames, London</famous-place></famous-place-group><famous-people-group><famous-people>East India Company (1660-1858)</famous-people><famous-people>Hawkins, Sir John (1532-1595)</famous-people><famous-people>Sharp, Granville (1735-1813)</famous-people><famous-people>Wheatley, Phillis (1753-1784)</famous-people><famous-people>Sancho, Ignatius (1729-1780)</famous-people><famous-people>Equiano, Olaudah (c.1745-1797)</famous-people><famous-people>Aldridge, Ira (1807-1867)</famous-people><famous-people>Seacole, Mary (1805-1881)</famous-people><famous-people>Lanier, Nicolas (1588-1666)</famous-people><famous-people>Papin, Denis (1647-1712)</famous-people><famous-people>Houblon, Sir John (1632-1712)</famous-people><famous-people>Garrick, David (1717-1779)</famous-people><famous-people>Swift, Jonathan (1667-1745)</famous-people><famous-people>Farquhar, George (1678-1707)</famous-people><famous-people>Woffington, Margaret (c 1720-1760) known as Peg Woffington</famous-people><famous-people>Goldsmith, Oliver (1728-1774)</famous-people><famous-people>Whittington, Richard (c 1350-1423)</famous-people></famous-people-group></subject-indexing><illustration-group><illustration><legal-status>Rights Reserved</legal-status></illustration></illustration-group><relations nbdc="1"><relation role="underResponsibilityOf" scheme="michael:/portal/relations/digital-collection" target="IS:40463"><abstract><p>London Metropolitan Archives</p></abstract><relation-infos type="institution"><institution id="IS:40463"><identification><name>London Metropolitan Archives</name><institution-type-group><institution-type code="t002" scheme="michael:/portal/institution-types"><fr>Archives</fr><en>Archive</en><cy>Archif</cy><it>Archivi</it><it>Archivio</it><cs>Archivy</cs><ee>Arhiiv</ee><fi>Arkisto</fi><sv>Arkiv</sv><el>Αρχείο</el><nl>Archief</nl><lv>Arhīvs</lv><hu>Archívum/Levéltár</hu><de>Archiv</de><bg>Архив</bg><sk>Archívy</sk><es>Archivo</es></institution-type></institution-type-group><administrative-status-group><administrative-status xml:lang="en">Public</administrative-status><administrative-status xml:lang="cy">Cyhoeddus</administrative-status><administrative-status xml:lang="it">Pubblico</administrative-status></administrative-status-group></identification><address><street-group><street>40 Northampton Road\n\n</street></street-group><locality>London</locality><postal-code>EC1R 0HB</postal-code><country code="gb" scheme="michael:/portal/countries"><cs>SPOJENÉ KRÁLOVSTVÍ (VELKÁ BRITÁNIE)</cs><ee>ÜHENDATUD KUNINGRIIGID</ee><en>UNITED KINGDOM</en><cy>Y DEYRNAS UNEDIG</cy><fr>ROYAUME UNI</fr><it>REGNO UNITO</it><fi>Iso-Britannia</fi><sv>STORBRITANNIEN</sv><el>ΗΝΩΜΕΝΟ ΒΑΣΙΛΕΙΟ</el><nl>VERENIGD KONINKRIJK</nl><lv>APVIENOTĀ KARALISTE</lv><de>VEREINIGTES KÖNIGREICH (UK)</de><pl>WIELKA BRYTANIA</pl><sk>Veľká Británia</sk><es>REINO UNIDO</es></country></address><communications><telephone>020 7332 3820</telephone><fax>020 7833 9136\n</fax><email-group><email>ask.lma@cityoflondon.gov.uk</email></email-group></communications><relations/><metadata><creation><date>2007-02-20T01:01:51CET</date><agent code="institution_server"/></creation><validation><date>2007-02-20T01:01:51CET</date><agent code="institution_server"/></validation><record-status code="valid"/><owner><agent code="institution_server"/></owner><language code="en"/></metadata></institution></relation-infos></relation><relation role="isMadeAvailableBy" scheme="michael:/portal/relations/digital-collection" target="UK-SE-e375541d"><relation-infos type="service"><service id="UK-SE-e375541d"><identification><title>London Metropolitan Archives: Learning Zone Website</title></identification><description><abstract><p>This website is aimed at children, teachers and life long learners. It provides public access to the resources of schooLMAte which explores history of London's communities with the help of Elma the Elf, through stories, images timelines, audio galleries, interactive maps and links. The site also provides access to the Black and Asian Londoners searchable database and to Archive Work which explains what happens in an archive and what is is like to work in one.</p></abstract><language-group><language code="eng" scheme="michael:/portal/languages"><cs>Angličtina</cs><ee>Inglise</ee><en>English</en><pl>English</pl><fr>anglais</fr><it>English</it><fi>englanti</fi><sv>engelska</sv><de>Englisch</de><el>Αγγλικά</el><nl>Engels</nl><hu>Angol</hu><bg>английски</bg><sk>anglicky</sk><es>inglés</es></language></language-group><maintenance>Live</maintenance><audience-group><audience code="au1" scheme="michael:/portal/audiences"><cs>Vědecký výzkum</cs><ee>Akadeemilised uuringud</ee><fr>Recherche</fr><en>Academic research</en><cy>Ymchwil academaidd</cy><it>Ricerca</it><el>Ακαδημαϊκή έρευνα</el><nl>Wetenschappelijk onderzoek</nl><lv>Akadēmiskiem pētījumiem</lv><hu>Kutatás</hu><de>Wissenschaftliche Forschung</de><pl>Poziom naukowy</pl><sk>Vedecký výskum</sk><es>Investigación académica</es><sv>Akademisk forskning</sv></audience><audience code="au2" scheme="michael:/portal/audiences"><cs>Děti</cs><ee>Lastele</ee><fr>Enfant</fr><en>Children</en><cy>Plant</cy><it>Infanzia</it><el>Παιδιά</el><nl>Kinderen</nl><lv>Bērniem</lv><hu>Gyermekek</hu><de>Kinder</de><pl>Dzieci</pl><sk>Deti</sk><es>Niños</es><sv>Barn</sv></audience><audience code="au3" scheme="michael:/portal/audiences"><cs>Studenti</cs><ee>Kooliharidus</ee><fr>Scolaire / étudiants</fr><en>Formal education</en><cy>Addysg ffurfiol</cy><it>Scuola</it><el>Επίσημη εκπαίδευση</el><nl>Onderwijs</nl><lv>Skolēniem</lv><hu>Oktatás</hu><de>Schulbildung</de><pl>Edukacja szkolna</pl><sk>Vzdelávanie</sk><es>Educación escolar</es><sv>Skola</sv></audience><audience code="au7" scheme="michael:/portal/audiences"><cs>Celoživotní vzdělávání</cs><ee>Elukestev õpe</ee><fr>Formation continue</fr><en>Life-long learning</en><cy>Dysgu gydol oes</cy><it>Formazione permanente</it><nl>Levenslang leren</nl><lv>Mūžizglītībai</lv><hu>Felnőttképzés</hu><de>Lebenslanges Lernen</de><el>Δια βίου μάθηση</el><pl>Edukacja ustawiczna</pl><sk>Celoživotné vzdelávanie</sk><es>Formación continua</es><sv>Livslångt lärande</sv></audience></audience-group><legal-status>Rights Reserved</legal-status></description><access><access-type-group><access-type code="at8" scheme="michael:/portal/access-types"><cs>Webová stránka</cs><sk>Webová stránka</sk><en>Website</en><nl>Website</nl><it>Sito web</it><fr>Site Internet</fr><ee>Veebisait</ee><cy>Gwefan</cy><fi>WWW-sivusto</fi><sv>Webbplats</sv><el>Διαδικτυακός τόπος</el><lv>Interneta lapa</lv><hu>Honlap</hu><de>Webangebot</de><pl>Strona internetowa</pl><es>Sitio web</es></access-type></access-type-group><wai code="w5" scheme="michael:/portal/wai"><cs>Netestováno</cs><ee>Pole testitud</ee><fr>Non testé</fr><de>Nicht geprüft</de><en>Not tested</en><it>Non testato</it><cy>Heb ei brofi</cy><fi>Ei testattu</fi><sv>Ej testad</sv><el>Μη δοκιμασμένο</el><nl>Niet getest</nl><lv>Nav testēts</lv><hu>Nem tesztelt</hu><pl>Niesprawdzone</pl><sk>netestované</sk><es>No testado</es></wai><technical-requirements><p>Adobe Flash is required for interactive maps</p></technical-requirements></access><access-locations><access-location><locator>http://www.corpoflondon.gov.uk/Corporation/lma_learning/</locator></access-location></access-locations><relations/><metadata><creation><date>2007-01-08T11:35:03</date><agent code="knorgrove" email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></creation><modification><date>2007-01-08T11:35:03</date><agent code="knorgrove" email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></modification><validation><date>2007-01-31T10:26:30</date><agent code="knorgrove" email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></validation><owner><agent code="knorgrove" email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></owner><update>2007-01-08</update><language code="en"/><record-status code="valid"/><rights>Licenced under the Creative Commons Licence
\n(http://creativecommons.org/licenses/by-nc-sa/2.0/uk/)</rights></metadata></service></relation-infos></relation><relation role="hasSource" scheme="michael:/portal/relations/digital-collection" target="cornucopia:6702"><abstract><p>History of London Collection</p></abstract></relation><relation role="isPartOf.inverse" target="UK-DC-efe68045"><relation-infos type="digital-collection"><digital-collection id="UK-DC-efe68045"><identification><title>London Metropolitan Archives: Black and Asian Londoners</title></identification><description><legal-status>Rights Reserved</legal-status><language-group><language code="eng" scheme="michael:/portal/languages"><cs>Angličtina</cs><ee>Inglise</ee><en>English</en><pl>English</pl><fr>anglais</fr><it>English</it><fi>englanti</fi><sv>engelska</sv><de>Englisch</de><el>Αγγλικά</el><nl>Engels</nl><hu>Angol</hu><bg>английски</bg><sk>anglicky</sk><es>inglés</es></language></language-group><abstract><p>This is a searchable database of over 2,000 entries for Black and Asian people, identified specifically as being of African or Asian orgin, through researching over 1,000 Anglican Baptism Registers in London. The earliest entry is from 1597 and the latest from 1856. The database is an important source of evidence for the fascinating presence, lives and experiences of Black and Asian people in London, spanning three centuries. The database is searchable by: surname; date; street; status; borough; and place or country of orgin.</p></abstract><digital-format-group><digital-format code="html" scheme="michael:/portal/digital-formats"><fr>HTML</fr><el>HTML</el><es>HTML</es><pl>HTML</pl><de>HTML</de><hu>HTML</hu><lv>HTML</lv><nl>HTML</nl><fi>HTML</fi><en>HTML</en><it>HTML</it><sv>HTML</sv><cy>HTML</cy><cs>HTML</cs><ee>HTML</ee><sk>html</sk></digital-format></digital-format-group><digital-type-group><digital-type code="dt001.1" scheme="michael:/portal/digital-types"><fr>Texte</fr><en>Text</en><cy>Testun</cy><it>Testo</it><cs>Text</cs><sk>Text</sk><sv>Text</sv><ee>Tekst</ee><pl>Tekst</pl><nl>Tekst</nl><fi>Tekstiaineistoa</fi><lv>Teksts</lv><hu>Szöveg</hu><el>Κείμενο</el><de>Text</de><es>Texto</es></digital-type><digital-type code="dt001.5" scheme="michael:/portal/digital-types"><en>Dataset</en><cy>Set ddata</cy><fr>Ensemble de données</fr><it>Dataset</it><sv>Dataset</sv><nl>Dataset</nl><cs>Soubor dat</cs><ee>Dataset</ee><fi>Tiedosto</fi><lv>Datu kopa</lv><hu>Adatbázis</hu><de>Datensatz</de><el>Σύνολο δεδομένων</el><pl>Kartoteka (zbiór danych)</pl><sk>Súbor údajov</sk><es>Conjunto de datos</es></digital-type></digital-type-group><size>c 2,000 entries</size><accrual>Complete</accrual><access-control>unrestricted</access-control></description><subject-indexing><subject-group><subject code="76" scheme="michael:/portal/subjects"><cs>Kulturní identita</cs><ee>Kultuuriline identiteet</ee><en>Cultural identity</en><it>Identità culturale</it><fr>Identité culturelle</fr><cy>Hunaniaeth ddiwylliannol</cy><fi>Kulttuuri-identiteetti</fi><sv>Kulturell identitet</sv><el>Πολιτιστική ταυτότητα</el><nl>Culturele identiteit</nl><lv>Kultūras identitāte</lv><hu>Kulturális identitás</hu><de>Kulturelle Identität</de><pl>Tożsamość kulturowa</pl><sk>Kultúrna identita</sk><es>Identidad cultural</es></subject><subject code="99" scheme="michael:/portal/subjects"><cs>Genealogie</cs><nl>Genealogie</nl><ee>Genealoogia</ee><en>Genealogy</en><it>Genealogia</it><fr>Généalogie</fr><cy>Achyddiaeth</cy><cy>Achyddiaeth (hanes teulu)</cy><fi>Sukututkimus</fi><sv>Genealogi</sv><el>Γενεαλογία</el><lv>Ģenealoģija</lv><hu>Genealógia (családtörténet)</hu><de>Genealogie</de><bg>Генеалогия</bg><pl>Genealogia</pl><sk>Genealógia</sk><es>Genealogía</es></subject><subject code="88" scheme="michael:/portal/subjects"><cs>Náboženství</cs><ee>Religioon</ee><en>Religion</en><it>Religione</it><fr>Religion</fr><sv>Religion</sv><cy>Crefydd</cy><fi>Uskonto</fi><el>Θρησκεία</el><nl>Religie</nl><lv>Reliģija</lv><hu>Vallás</hu><de>Religion</de><bg>Религия</bg><pl>Religia</pl><sk>Náboženstvo</sk><es>Religión</es></subject><subject code="207" scheme="michael:/portal/subjects"><cs>Obyvatelstvo</cs><ee>Rahvastik</ee><en>Population</en><it>Etnografia</it><fr>Population</fr><cy>Poblogaeth</cy><fi>Kansatiede</fi><sv>Befolkning</sv><el>Πληθυσμός</el><nl>Bevolking</nl><lv>Iedzīvotāji</lv><hu>Népesség</hu><de>Bevölkerung</de><bg>Население</bg><pl>Populacja</pl><sk>Obyvateľstvo</sk><es>Población</es></subject><subject code="212" scheme="michael:/portal/subjects"><cs>Etnické menšiny</cs><sk>Etnické menšiny</sk><ee>Mustad ja etnilised vähemused</ee><en>Black and ethnic minority groups</en><it>Minoranze etniche</it><fr>Minorité ethnique</fr><cy>Grwpiau pobl dduon a lleiafrifoedd ethnig</cy><fi>Vähemmistöt</fi><sv>Färgade och etniska minoriteter</sv><el>Μειονοτικές και εθνοτικές ομάδες</el><nl>Etnische minderheden</nl><lv>Etnisko minoritāšu grupas</lv><hu>Etnikumok, kisebbségek</hu><de>Ethnische Minderheiten</de><bg>Малцинства</bg><pl>Mniejszości etniczne</pl><es>Grupos minoritarios étnicos</es></subject><subject code="290" scheme="michael:/portal/subjects"><cs>Pracovní prostředí</cs><ee>Töökeskkond</ee><en>Work environment</en><it>Ambiente di lavoro</it><fr>Milieu du travail</fr><fi>Työympäristö</fi><sv>Arbetsmiljö</sv><el>Περιβάλλον εργασίας</el><nl>Arbeidsmilieu</nl><lv>Darba vide</lv><hu>Munkakörnyezet</hu><de>Arbeitsleben</de><bg>Работна среда</bg><pl>Środowisko pracy</pl><sk>Pracovné prostredie</sk><es>Ambiente de trabajo</es></subject><subject code="225" scheme="michael:/portal/subjects"><cs>Rozvoj měst</cs><ee>Linna-areng</ee><en>Urban development</en><it>Sviluppo urbano</it><fr>Développement urbain</fr><cy>Datblygu trefol</cy><fi>Kaupunkisuunnittelu</fi><sv>Urban utveckling</sv><el>Αστική ανάπτυξη</el><nl>Stedelijke ontwikkeling</nl><lv>Pilsētu attīstība</lv><hu>Városfejlődés</hu><de>Stadtentwicklung</de><bg>Урбанистично развитие</bg><pl>Rozwój miast</pl><sk>Rozvoj miest</sk><es>Desarrollo urbano</es></subject></subject-group><period-group><period

```
code="c16" scheme="michael:/portal/periods"><ee>16. sajand</ee><cs>16. století</cs><fr>16e siècle</fr><en>16th
century</en><it>XVI secolo</it><fi>1500 -luku</fi><sv>1500-talet</sv><el>16ος αιώνας</el><nl>16e eeuw</nl><lv>16.
gadsimts</lv><hu>16. század</hu><de>16. Jahrhundert</de><bg>16 век</bg><cy>16eg ganrif</cy><pl>XVI
wiek</pl><sk>16.storočie</sk><es>Siglo XVI</es></period><period code="c17"
scheme="michael:/portal/periods"><ee>17. sajand</ee><cs>17. století</cs><fr>17e siècle</fr><en>17th
century</en><it>XVII secolo</it><fi>1600 -luku</fi><sv>1600-talet</sv><el>17ος αιώνας</el><nl>17e eeuw</nl><lv>17.
gadsimts</lv><hu>17. század</hu><de>17. Jahrhundert</de><bg>17 век</bg><cy>17eg ganrif</cy><pl>XVII
wiek</pl><sk>17.storočie</sk><es>Siglo XVII</es></period><period code="c18"
scheme="michael:/portal/periods"><ee>18. sajand</ee><cs>18. století</cs><fr>18e siècle</fr><en>18th
century</en><it>XVIII secolo</it><fi>1700 -luku</fi><sv>1700-talet</sv><el>18ος αιώνας</el><nl>18e
eeuw</nl><lv>18. gadsimts</lv><hu>18. század</hu><de>18. Jahrhundert</de><bg>18 век</bg><cy>18fed
ganrif</cy><pl>XVIII wiek</pl><sk>18.storočie</sk><es>Siglo XVIII</es></period><period code="c19"
scheme="michael:/portal/periods"><ee>19. sajand</ee><cs>19. století</cs><fr>19e siècle</fr><en>19th
century</en><it>XIX secolo</it><fi>1800 -luku</fi><sv>1800-talet</sv><el>19ος αιώνας</el><nl>19e eeuw</nl><lv>19.
gadsimts</lv><hu>19. század</hu><de>19. Jahrhundert</de><bg>19 век</bg><cy>19eg ganrif</cy><pl>XIX
wiek</pl><sk>19.storočie</sk><es>Siglo XIX</es></period></period-group><culture-group><culture code="afr"
scheme="michael:/portal/civilisations"><cs>Africká</cs><en>African</en><es>Africa</es><cy>Affricanaidd</cy><fi>Afr
ikkalainen</fi><sv>Afrikansk</sv><nl>Afrikaanse</nl><hu>Afrikai</hu><pl>Afrykańska</pl><sv>Afrikansk</sv></culture
><culture code="car"
scheme="michael:/portal/civilisations"><cs>Karibská</cs><en>Caribbean</en><es>Caribe</es><cy>Caribî
©dd</cy><fi>Karibialainen</fi><sv>Karibisk</sv><nl>Caraïbische</nl><hu>Karibi</hu><pl>Karaibska</pl><sv>Karibisk</
sv></culture><culture code="sas" scheme="michael:/portal/civilisations"><cs>Jihoasijská</cs><en>South
Asian</en><es>Asia Meridional</es><cy>De Asiaidd</cy><fi>Etelä-Aasialainen</fi><sv>Sydasiatisk</sv><nl>Zuid-
Aziatische</nl><hu>Dél-ázsiai</hu><pl>Południowoazjatycka</pl><sv>Sydasiatisk</sv></culture><culture code="weu"
scheme="michael:/portal/civilisations"><cs>Západoevropská</cs><en>Western European</en><es>Europa
Occidental</es><cy>Gorllewin Ewrop</cy><fi>Länsi-Eurooppalainen</fi><sv>Västeuropeisk</sv><nl>West-
Europese</nl><hu>Nyugat-európai</hu><pl>Zachodnioeuropejska</pl><sv>Västeuropeisk</sv></culture></culture-
group><spatial-coverage-group><spatial-coverage code="gb" scheme="michael:/portal/spatial-coverages"><cs>SPOJENÉ
KRÁLOVSTVÍ (VELKÁ BRITÁNIE)</cs><ee>Ühendkuningriik</ee><en>UNITED KINGDOM</en><cy>Y DEYRNAS UNEDIG</cy><it>REGNO
UNITO</it><fr>ROYAUME UNI</fr><fi>Iso-Britannia</fi><sv>Storbrittanien</sv><el>ΗΝΩΜΕΝΟ ΒΑΣΙΛΕΙΟ</el><nl>VERENIGD
KONINKRIJK</nl><lv>APVIENOTĀ KARALISTE</lv><de>VEREINIGTES KÖNIGREICH (UK)</de><sk>Spojené
kráľovstvo</sk><es>REINO UNIDO</es></spatial-coverage><spatial-coverage code="gbe"
scheme="michael:/portal/spatial-
coverages"><cs>Anglie</cs><ee>Inglismaa</ee><en>England</en><it>Inghilterra</it><fr>Angleterre</fr><fi>Englanti</f
i><sv>England</sv><nl>Engeland</nl><lv>Anglija</lv><cy>Lloegr</cy><sk>Anglicko</sk><es>Inglaterra</es></spatial-
coverage><spatial-coverage code="in" scheme="michael:/portal/spatial-
coverages"><cs>INDIE</cs><ee>India</ee><sk>India</sk><en>INDIA</en><cy>INDIA</cy><it>INDIA</it><fr>INDE</fr><fi>In
tia</fi><sv>Indien</sv><el>ΙΝΔΙΑ</el><nl>INDIA</nl><es>INDIA</es><lv>INDIJA</lv><de>INDIEN</de></spatial-
coverage><spatial-coverage code="w0019" scheme="michael:/portal/spatial-coverages"><cs>Severní
Afrika</cs><ee>Põhja-Aafrika</ee><en>North Africa</en><en>North Africa</en><it>Africa
settentrionale</it><fr>Afrique du Nord</fr><fi>Pohjois-Afrikka</fi><sv>Nordafrika</sv><cy>Gogledd
Affrica</cy><de>Nordafrika</de><el>Βόρεια Αφρική</el><nl>Noord Afrika</nl><lv>Ziemeļāfrika</lv><sk>Severná
Afrika</sk><es>África septentrional</es></spatial-coverage><spatial-coverage code="w0022"
scheme="michael:/portal/spatial-coverages"><cs>Jižní Afrika</cs><ee>Lõuna-Aafrika</ee><en>Southern
Africa</en><it>Africa meridionale</it><fr>Afrique australe</fr><fi>Eteläinen Afrikka</fi><sv>Södra
Afrika</sv><el>Νότια Αφρική</el><nl>Zuid Afrika</nl><lv>Dienvidāfrika</lv><de>Südliches Afrika</de><sk>Južnejšia
Afrika</sk><es>África meridional</es></spatial-coverage><spatial-coverage code="w0023"
scheme="michael:/portal/spatial-coverages"><cs>Střední Afrika</cs><ee>Kesk-Aafrika</ee><en>Central
Africa</en><it>Africa centrale</it><fr>Afrique centrale</fr><fi>Keski-
Afrikka</fi><sv>Centralafrika</sv><en>Κεντρική Αφρική</en><nl>Centraal
Afrika</nl><lv>Centrālāfrika</lv><de>Zentralafrika</de><sk>Stredná Afrika</sk><es>África central</es></spatial-
coverage><spatial-coverage code="w0025" scheme="michael:/portal/spatial-coverages"><cs>Západní
Afrika</cs><ee>Lääne-Aafrika</ee><en>West Africa</en><it>Africa occidentale</it><fr>Afrique de
l'Ouest</fr><fi>Länsi-Afrikka</fi><sv>Västafrika</sv><el>Δυτική Αφρική</el><nl>West
Afrika</nl><lv>Rietumāfrika</lv><de>Westafrika</de><sk>Západná Afrika</sk><es>África occidental</es></spatial-
coverage></spatial-coverage-group><start-date>1597</start-date><end-date>1856</end-date></subject-
indexing><illustration-group><illustration><legal-status>Rights Reserved</legal-
status></illustration></illustration-group><relations/><metadata><creation><date>2007-01-08T12:00:33</date><agent
code="knorgrove" email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></creation><modification><date>2007-
01-31T09:54:49</date><agent code="knorgrove" email="katie.norgrove@hotmail.co.uk" name="Katie
Norgrove"/></modification><validation><date>2007-01-31T10:02:23</date><agent code="knorgrove"
email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></validation><owner><agent code="knorgrove"
email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></owner><update>2007-01-08</update><language
code="en"/><record-status code="valid"/><rights>Licenced under the Creative Commons
Licence&#13;\n(http://creativecommons.org/licenses/by-nc-sa/2.0/uk/)</rights></metadata></digital-
collection></relation-infos></relation></relations><metadata><creation><date>2007-01-04T11:46:27</date><agent
code="knorgrove" email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></creation><modification><date>2008-
02-10T20:29:28</date><agent code="kfernie" email="kate.fernie@mla.gov.uk" name="Kate
Fernie"/></modification><validation><date>2007-01-31T10:03:16</date><agent code="knorgrove"
email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></validation><owner><agent code="knorgrove"
email="katie.norgrove@hotmail.co.uk" name="Katie Norgrove"/></owner><update>2007-01-04</update><language
code="en"/><record-status code="valid"/><rights>Licenced under the Creative Commons
Licence&#13;\n(http://creativecommons.org/licenses/by-nc-
sa/2.0/uk/)</rights><source>UK</source></metadata></digital-collection>
```


9.2 RDF File

```
<rdf:RDF
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:foaf="http://xmlns.com/foaf/0.1/"
```

```
    xmlns:baseURI="http://mint.image.ece.ntua.gr/resource/"
    xmlns:taxon="http://lod.taxonconcept.org/ontology/txn.owl#"
    xmlns:relationship="http://purl.org/vocab/relationship/"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:dcterms="http://purl.org/dc/terms/"
    xmlns:olo="http://purl.org/ontology/olo/core#"
    xmlns:co="http://purl.org/ontology/co/core#"
    xmlns:j.0="http://purl.org/dc/dcmitype/"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >
  <rdf:Description rdf:about="http://mint.image.ece.ntua.gr/resource/UK-DC-7036a5f">
    <co:event>Great Plague of London, 1665</co:event>
    <dcterms:coverage>River Thames, London</dcterms:coverage>
    <relationship:participant>Garrick, David (1717-1779)</relationship:participant>
    <dc:subject>Religion</dc:subject>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/John_Houblon"/>
    <relationship:participant>East India Company (1660-1858)</relationship:participant>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/John_Hawkins"/>
    <dc:date>18th century</dc:date>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/East_India_Company"/>
    <dcterms:language rdf:resource="http://dbpedia.org/resource/French_language"/>
    <dc:subject>Leisure</dc:subject>
    <dcterms:coverage>Guildhall, London</dcterms:coverage>
    <relationship:participant>Goldsmith, Oliver (1728-1774)</relationship:participant>
    <rdf:type rdf:resource="http://purl.org/dc/dcmitype/Collection"/>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/Mary_Seacole"/>
    <dc:subject>Work environment</dc:subject>
    <dcterms:spatial>England</dcterms:spatial>
    <relationship:participant>Swift, Jonathan (1667-1745)</relationship:participant>
    <dc:subject>Legal history</dc:subject>
    <dcterms:rights>Rights Reserved</dcterms:rights>
    <co:event>Great Fire of London, 1666</co:event>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/Ira_Aldridge"/>
    <co:event>Black Death, 1348</co:event>
    <dcterms:coverage>Westminster Bridge, London</dcterms:coverage>
    <relationship:participant>Whittington, Richard (c 1350-1423)</relationship:participant>
    <relationship:participant>Sancho, Ignatius (1729-1780)</relationship:participant>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/Denis_Papin"/>
    <dcterms:format>Macromedia Flash</dcterms:format>
    <dcterms:coverage>Monument, London</dcterms:coverage>
    <dcterms:spatial rdf:resource="http://dbpedia.org/resource/England"/>
    <dc:date>21st century</dc:date>
    <dc:title>London Metropolitan Archives: schooLMAte</dc:title>
    <dcterms:format>JPEG</dcterms:format>
    <dc:subject>Migration</dc:subject>
    <relationship:participant>Papin, Denis (1647-1712)</relationship:participant>
    <relationship:participant>Hawkins, Sir John (1532-1595)</relationship:participant>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/Granville_Sharp"/>
    <dc:subject>History</dc:subject>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/Oliver_Goldsmith"/>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/Ignatius_Sancho"/>
    <dc:date>17th century</dc:date>
    <dc:subject>Transport</dc:subject>
    <dcterms:format>HTML</dcterms:format>
    <dc:subject>Graphic arts</dc:subject>
    <dc:subject>Urban development</dc:subject>
    <dcterms:language rdf:resource="http://dbpedia.org/resource/English_language"/>
    <dc:date>Medieval</dc:date>
    <relationship:participant>Farquhar, George (1678-1707)</relationship:participant>
    <taxon:endDate>2003</taxon:endDate>
    <dc:subject>Buildings and the build environment</dc:subject>
    <relationship:participant rdf:resource="http://dbpedia.org/resource/Phillis_Wheatley"/>
    <dc:date>High and late medieval</dc:date>
    <relationship:participant>Equiano, Olaudah (c.1745-1797)</relationship:participant>
    <dc:date>19th century</dc:date>
    <dc:date>16th century</dc:date>
    <taxon:startDate>1067</taxon:startDate>
    <dcterms:coverage>Old London Bridge</dcterms:coverage>
    <dc:subject>Social policy and welfare</dc:subject>
    <dcterms:spatial>UNITED KINGDOM</dcterms:spatial>
    <dc:date>20th century</dc:date>
    <dc:subject>Buildings and the built environment</dc:subject>
    <relationship:participant>Aldridge, Ira (1807-1867)</relationship:participant>
    <relationship:participant>Woffington, Margaret (c 1720-1760) known as Peg
Woffington</relationship:participant>
```

```
    &lt;dcterms:language&gt;English&lt;/dcterms:language&gt;
    &lt;dc:subject&gt;Black and ethnic minority groups&lt;/dc:subject&gt;
    &lt;relationship:participant&gt;Lanier, Nicolas (1588-1666)&lt;/relationship:participant&gt;
    &lt;dc:subject&gt;Cultural identity&lt;/dc:subject&gt;
    &lt;relationship:participant&gt;Sharp, Granville (1735-1813)&lt;/relationship:participant&gt;
    &lt;dc:identifier&gt;UK-DC-7036a5f&lt;/dc:identifier&gt;
    &lt;relationship:participant rdf:resource="http://dbpedia.org/resource/Richard_Whittington"/&gt;
    &lt;relationship:participant&gt;Houblon, Sir John (1632-1712)&lt;/relationship:participant&gt;
    &lt;relationship:participant rdf:resource="http://dbpedia.org/resource/George_Farquhar"/&gt;
    &lt;relationship:participant rdf:resource="http://dbpedia.org/resource/Jonathan_Swift"/&gt;
    &lt;dc:subject&gt;Health &amp; lifestyle&lt;/dc:subject&gt;
    &lt;dcterms:format&gt;Real Audio&lt;/dcterms:format&gt;
    &lt;relationship:participant rdf:resource="http://dbpedia.org/resource/David_Garrick"/&gt;
    &lt;relationship:participant rdf:resource="http://dbpedia.org/resource/Olaudah_Equiano"/&gt;
    &lt;relationship:participant&gt;Wheatley, Phillis (1753-1784)&lt;/relationship:participant&gt;
    &lt;dcterms:language&gt;French&lt;/dcterms:language&gt;
    &lt;dc:type&gt;Western European&lt;/dc:type&gt;
    &lt;dcterms:spatial rdf:resource="http://dbpedia.org/resource/United_Kingdom"/&gt;
    &lt;relationship:participant rdf:resource="http://dbpedia.org/resource/Margaret_Woffington"/&gt;
    &lt;relationship:participant&gt;Seacole, Mary (1805-1881)&lt;/relationship:participant&gt;
    &lt;dc:description&gt;This collection of digital images, interpretative text, maps and audioclips
allows children and teachers to explore the complex histories of: Black and Asian Londoners; the
French community; the Irish community; and the communities of the City of London.  The resources
draw on important photographs, designs, letters, prints, maps and interviews to illustrate the
important contribution that these communities have made to the making of London from 1067 to
2003.&lt;/dc:description&gt;
    &lt;dc:subject&gt;Civil, military and mining engineering&lt;/dc:subject&gt;
    &lt;dc:subject&gt;Politics and government&lt;/dc:subject&gt;
  &lt;/rdf:Description&gt;
  &lt;rdf:Description rdf:about="http://mint.image.ece.ntua.gr/resource/UK-DC-7036a5f.rdf"&gt;
    &lt;dcterms:date rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime"&gt;2011-11-
13T17:20:21.626Z&lt;/dcterms:date&gt;
    &lt;dcterms:publisher rdf:resource="http://www.image.ece.ntua.gr"/&gt;
    &lt;dcterms:creator rdf:resource="http://www.image.ece.ntua.gr/~nsimou/"/&gt;
    &lt;rdfs:label xml:lang="en"&gt;Document describing the Michael item having ID = UK-DC-
7036a5f&lt;/rdfs:label&gt;
    &lt;rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Document"/&gt;
  &lt;/rdf:Description&gt;
&lt;/rdf:RDF&gt;
```