Research Libraries UK Linked Open Data at The European Library

Introduction
A subset of the RLUK Database is available through The European Library data portal, published as Linked Open Data and searchable via an API. This document describes this subset of the RLUK Database, and the access mechanisms provided for the Hackathon participants to interact with the dataset.

The RLUK Database subset
The RLUK Database is an established Union Catalogue provided by Research Libraries UK. Hackathon participants will be able to interact with a subset of the RLUK database comprising nearly 20 million bibliographic records from 34 libraries.

The dataset is published in RDF, and linked to external datasets available as Linked Open Data. The major external datasets linked to are the following:

- VIAF Virtual Union Authority File
- Geonames
- Library of Congress Subject Headings
- Library of Congress Children’s Subject Headings
- Library of Congress Classification
- data.bnf.fr
- Gemeinsame Normdatei
- Dewey Decimal Classification
- ISO639-2 Languages
- MARC Countries

Statistics regarding the links to external datasets from the RLUK dataset can be found in the document: “Research Libraries UK Linked Open Data at The European Library: Dataset Content Statistics”. 
Obtaining the dataset

The dataset is available for download. A complete data dump of the dataset is available in two RDF representations: RDF/XML Syntax, and Turtle - Terse RDF Triple Language. Smaller samples of the data are also provided to provide hackathon participants with a practical way of starting to explore the dataset.

The data can be obtained in the following locations and formats:

- **Complete dataset files**
  - In Turtle (text/turtle) - a1195-complete.ttl.gz (6.1 GB)
  - In RDF/XML (application/rdf+xml) - a1195-complete.rdf.gz (6.9 GB)

- **Samples of 1 million records**
  - In Turtle (text/turtle) - a1195-sample-1million.ttl.gz (335 MB)
  - In RDF/XML (application/rdf+xml) - a1195-sample-1million.rdf.gz (381 MB)

- **Samples of 1 thousand records**
  - In Turtle (text/turtle) - a1195-sample-1thousand.ttl.gz (349 KB)
  - In RDF/XML (application/rdf+xml) - a1195-sample-1thousand.rdf.gz (387 KB)

Interacting with the dataset

In addition to the dataset exports available, hackathon participants will be able to interact online with the dataset. Two mechanisms are provided that allow searching and direct access to the bibliographic data by the resources URIs.

W3C Linked Data Platform access

The portal and data.theeuropeanlibrary.org both support LOD Content Negotiation according to W3C Linked Data Platform 1.0.

Direct access to the data of individual bibliographic records can be obtained by accessing the bibliographic records URIs found in the dataset exports. The access responses to URIs is done according to LOD Content Negotiation, as follows:

- Programmatic access to the URIs, or through data browsers, should specify the RDF representation desired by specifying an HTTP header “Accept”. The following RDF representations (mime types) are available:
  - text/turtle
Content negotiation is supported in data.theeuropeanlibrary.org. Any request to the URIs made with a web browser will be redirected to the corresponding HTML page at the portal. The web browser requests for the URIs are sent with the HTTP header `Accept:text/html` or `Accept:application/xhtml+xml`, therefore they are redirected to the corresponding HTML page at the portal.

Content negotiation is also available in the portal. An RDF record may be requested for the pages that present individual bibliographic resources. Any request to these pages with the HTTP header `Accept:text/turtle`, or other RDF specific mime types, will be redirected to the corresponding RDF record at data.theeuropeanlibrary.org.

**OpenSearch API**

The European Library provides an API that enables software developers to access the RLUK dataset.

The European Library API consists of an OpenSearch API, providing the faceted searches on languages, contributors, disciplines, countries and collections. The API returns results as raw data, in machine-friendly formats, such as JSON and XML.

Instructions on how to access the API, as well as its detailed technical documentation is available in the document: “The European Library API V2.0”.

www.theeuropeanlibrary.org

National Library of the Netherlands, NL 2509LK, The Hague
Data model and vocabularies

The RLUK dataset is published under a data model that makes use of the RDA Element Vocabularies as a basis. Additional vocabularies are used for data elements whose semantics could not be expressed using the RDA Element Vocabularies. The data is represented using classes and properties defined in the following vocabularies:

- RDA Classes - http://rdaregistry.info/Elements/c/
- RDA Agent properties - http://rdaregistry.info/Elements/a/
- RDA Manifestation properties - http://rdaregistry.info/Elements/m/
- RDA Unconstrained properties - http://rdaregistry.info/Elements/u/
- RDA Work properties - http://rdaregistry.info/Elements/w/
- RDA Content Type vocabulary - http://rdvocab.info/termList/RDAContentType
- RDA Mode of Issuance vocabulary - http://rdvocab.info/termList/ModelIssue
- RDA Illustrative Content vocabulary - http://rdvocab.info/termList/IllusContent
- FRBRer model - http://iflastandards.info/ns/fr/frbr/frbrer/
- Europeana Data Model - http://www.europeana.eu/schemas/edm/
- OWL 2 Web Ontology Language - http://www.w3.org/2002/07/owl#
- Dublin Core Terms - http://purl.org/dc/terms/
- Metadata Object Description Schema - http://www.loc.gov/mods/rdf/v1#
- WGS84 Geo Positioning - http://www.w3.org/2003/01/geo/wgs84_pos#
- OAI ORE Terms - http://www.openarchives.org/ore/terms/
- RDF Schema vocabulary - http://www.w3.org/2000/01/rdf-schema#

A complete description of the data model used for the RLUK dataset can be found in the document: “Linked Data at The European Library: Data Model and Vocabularies”.

Statistics regarding the usage of data elements in the RLUK dataset can be found in the document: “Research Libraries UK Linked Open Data at The European Library: Dataset Content Statistics”.
Other relevant data sets

There are other data sets which may be of interest to those using RLUK and TEL linked data. A few of these are described briefly here with links to further documentation where available.

### The British National Bibliography

The British National Bibliography (BNB) contains details of publications printed in the United Kingdom and the Republic of Ireland since 1950, and electronic publications since 2003. The BNB is published as linked and open data (under a CC0 public domain dedication licence) and can be accessed in several ways:

- **Dataset access**
  - Simple search [http://bnb.data.bl.uk](http://bnb.data.bl.uk)
  - SPARQL Endpoint [http://bnb.data.bl.uk/flint](http://bnb.data.bl.uk/flint)
  - Full monthly exports in RDF/XML and N-triples [http://www.bl.uk/bibliographic/download.html#lodbnb](http://www.bl.uk/bibliographic/download.html#lodbnb)
    - Includes files containing subsets of BNB, a more basic RDF expression of BNB and of further books not eligible for inclusion in BNB
  - Weekly additions to BNB in RDF/XML [http://www.bl.uk/bibliographic/bnbrdfxml.html](http://www.bl.uk/bibliographic/bnbrdfxml.html)

- **Vocabularies**
  - Bio 0.1, a vocabulary for publishing biographical information - [http://purl.org/vocab/bio/0.1/](http://purl.org/vocab/bio/0.1/)
  - British Library Terms, a small vocabulary complementing existing bibliographic schemas - [http://www.bl.uk/schemas/bibliographic/blterms](http://www.bl.uk/schemas/bibliographic/blterms)
  - Dublin Core - [http://purl.org/dc/terms/](http://purl.org/dc/terms/)
  - Event Ontology - [http://purl.org/NET/c4dm/event.owl](http://purl.org/NET/c4dm/event.owl)
  - FOAF - [http://xmlns.com/foaf/0.1/](http://xmlns.com/foaf/0.1/)
  - ISBD, a vocabulary from the International Standard Bibliographic Description standard - [http://iflastandards.info/ns/isbd/elements/](http://iflastandards.info/ns/isbd/elements/)
  - Org, terms supporting linked data publishing of organizational information - [http://www.w3.org/ns/org](http://www.w3.org/ns/org)
  - OWL 2 Web Ontology Language - [http://www.w3.org/2002/07/owl#](http://www.w3.org/2002/07/owl#)
  - SKOS, terms supporting the publication of subject classifications and taxonomies - [http://www.w3.org/2004/02/skos/](http://www.w3.org/2004/02/skos/)
  - WGS84 Geo Positioning - [http://www.w3.org/2003/01/geo/wgs84_pos#](http://www.w3.org/2003/01/geo/wgs84_pos#)
  - RDA - [http://rdaregistry.info/Elements](http://rdaregistry.info/Elements)
Full documentation on the data model, schema and URI patterns is available at [http://bnb.data.bl.uk/docs](http://bnb.data.bl.uk/docs)

**The British Museum Semantic Web Collection Online**

The British Museum Semantic Web Collection provides access to data about approximately 3,500,000 objects in the British Museum’s collections. The data is organised using the CIDOC CRM (Conceptual Reference Model) (more information at [http://www.cidoc-crm.org](http://www.cidoc-crm.org)).

- **Dataset access**
  - HTML web search/result for collections [http://www.britishmuseum.org/research/collection_online/collection_search_results.aspx](http://www.britishmuseum.org/research/collection_online/collection_search_results.aspx)
  - SPARQL Endpoint [http://collection.britishmuseum.org/sparql](http://collection.britishmuseum.org/sparql)
  - Download (NQuads) [http://collection.britishmuseum.org/dumps/](http://collection.britishmuseum.org/dumps/) (1.6 GB)

- **Vocabularies**
  - SKOS, terms supporting the publication of subject classifications and taxonomies - [http://www.w3.org/2004/02/skos/](http://www.w3.org/2004/02/skos/)

- **Licensing information**
  - [http://collection.britishmuseum.org/licensing.html](http://collection.britishmuseum.org/licensing.html)

- **URI Patterns and general help**

Further information is available at [http://collection.britishmuseum.org](http://collection.britishmuseum.org)

**Contact**

We welcome feedback and are particularly keen to hear about requests for improvement, as this will guide our future work. Questions and requests about linked open data at The European Library may be sent by email to info@theeuropeanlibrary.org, or on twitter @EuropeanLibrary.