

Case study: The British National Bibliography as Linked Open Data

Alan Danskin

Metadata Standards Manager



UK's National Library

- Reference Services
- Document Supply
- Legal Deposit of UK Publications
- Bibliographic Services



Boston Spa, Yorkshire

>150 million items

- Printed text
- MSS & archives
- Music scores
- Recordings Images
- Patents
- Philatelic items

Why Metadata is Important to the BL

Without adequate, accessible metadata:

- Our collection is effectively invisible
- We could not respond adequately to evolving user expectations
- We could not offer bibliographic services to the wider community



Definitions

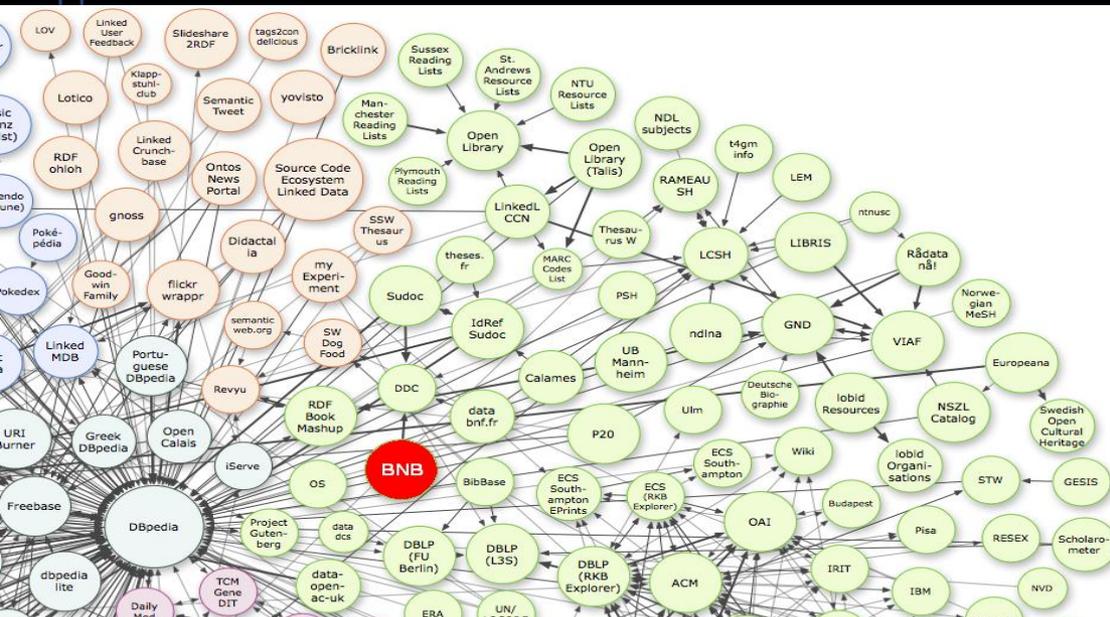


"A piece of data is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike."

<http://opendefinition.org/>

Linked Data

- Use **URIs** to identify things
- Use **HTTP URIs** so that these things can be referred to and looked up (**dereferenced**) by people and user agents
- Provide useful information about the thing when its URI is dereferenced, using standard formats such as **RDF/XML**
- Include links to other, related URIs in the exposed data to improve discovery of other related information on the Web



Sir Tim Berners-Lee

A Changing Environment

- The Web helped create a **collaboration culture & expectation that information should be open to all**
- Data has gained new value - **from new technologies enabling easy dissemination, connection & reuse**
- Knowledge organisations **must respond to expectations - increasing community relevance, stimulating innovation & knowledge creation**
- Open metadata is an **evolutionary step for the library principle of freedom of access to information**

The collage features several key elements:

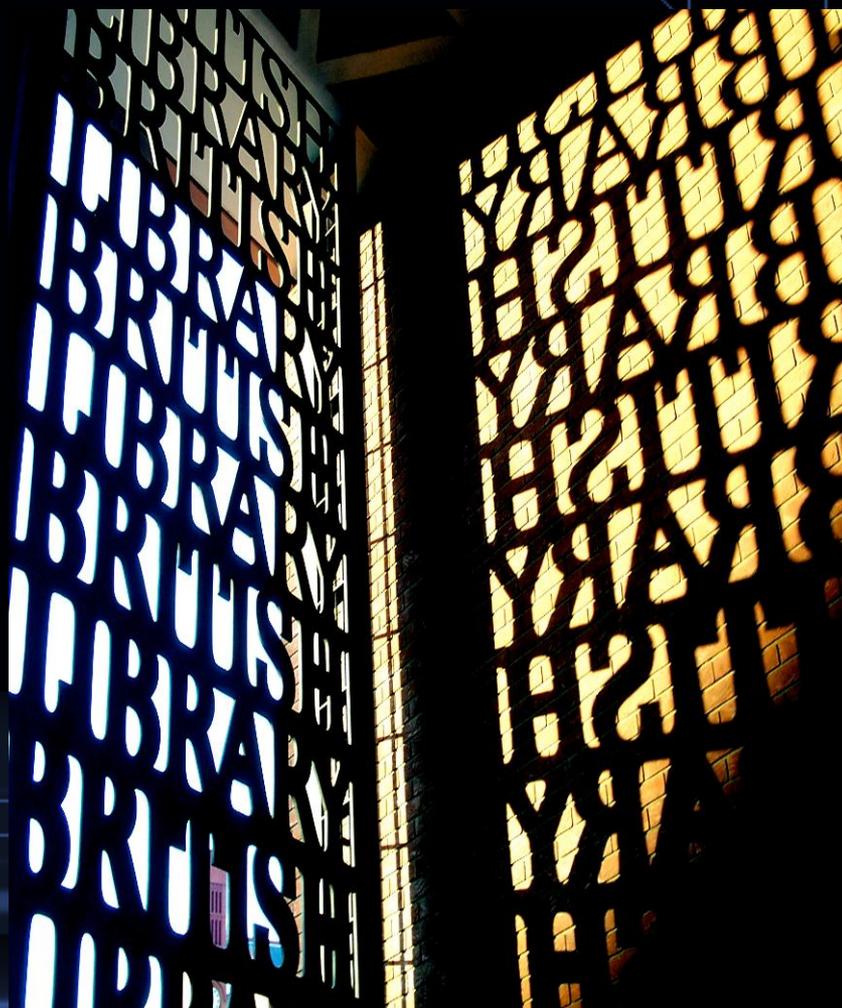
- EU Calendar:** A screenshot of the European Commission's website with the headline "The European Commission proposes an open data strategy".
- data.gov.uk:** A screenshot of the UK's public data portal with the slogan "Opening up government".
- Cabinet Office:** A screenshot of the UK Cabinet Office website with the headline "Public Data Corporation to free up public data and drive innovation".
- Open Data Case Studies:** A screenshot of a blog post titled "Open Data Case Studies" featuring a photo of Charlotte Aldritt and the text "Posted by Charlotte Aldritt on 27/06/2012".
- Open Data White Paper:** A screenshot of the "Open Data White Paper: Unleashing the Potential" document cover, featuring the Royal Coat of Arms and the text "Presented to Parliament by the Minister of State for the Cabinet Office and Minister General in Command of Her Majesty's June 2012".

BL Open Metadata Strategy

To meet these challenges the BL developed an open metadata strategy freely offering foundational metadata to the community

Objectives:

- **Collaboration** *with communities on innovative new services advancing debate from theory to practice*
- **Remove barriers** *to enable innovation without unnecessary restrictions*
- **Recognition** *while adopting licensing appropriate to the widest re-use (CC0)*



Objectives

Our strategy addressed the varying needs of:

- Traditional libraries
- Researchers using new metadata processing techniques
- New linked data developers

And offered opportunities for:

- Collaborations with open data communities, commercial partners & individuals
- Staff and organisational development & innovation using new technologies
- Improved web integration of library resources increasing visibility & reaching new users



Home > Metadata Services > Data Services > Sample Data

Sample Data

Free Data

The following sample files are currently available to download:

BNB in RDF/XML

This file provides a representation of a sample of the British National Bibliography in RDF/XML. Data is described using properties from the following vocabularies:

Dublin Core

SKOS

OWL, Time Ontology in OWL

RDF Schema

Bibliographic Ontology

and the preliminary versions of the **ISBD** and **RDA**

[[bnbrdfxml](#)] [1.5MB ZIP Format]

[Version 0.3.2, 21 January 2011]

Changes from the previous version are documented in [this table](#) (PDF format 55KB).

BNB in RDF/XML with links to external data sources

This file is an experimental file; it is a representation of a sample of the British National Bibliography in RDF/XML with links to the following linked data sources:

LCSH

MARC country and language codes

Dewey.info

Sections:

[British National Bibliography](#)

[Data Services](#)

> [Free data](#)

> [Priced data](#)

> [Licensing arrangements](#)

> [Technical information](#)

> [Sample data](#)

[Standards](#)

[News](#)

[Contact us](#)

This page contains links to Adobe PDF files. Accessibility solutions and free 'Reader' software are available from Adobe.

Open Data Challenges

Benefits of open data are significant but financial, reputational & legal challenges can exist *e.g. how to:*

- Capture the value of free offerings & gain recognition for resources invested
- Ensure community & Government expectations are met *e.g. via rival interpretations of 'free' & 'open'*
- License derived 3rd party data to support open re-supply
- Manage revenue & branding issues



Open Metadata Strategy

Services and products



Traditional Users Libraries / Bibliographic Agencies	“Researchers / Developers“	Linked Data Community
FREE Z39.50 Registration download	FREE FTP Registration required for monthly update	FREE SPARQL Triple dump
CHARGED MARC Exchange files		
MARC 21	Dublin Core RDF/XML JSON	LOD RDF

Our Linked Data Journey

What to Offer?

We wanted to:

- **Advance debate** from theory to practice via release of a 'critical mass' of data
- **Show commitment** by using a large, core dataset: niche examples are not as compelling

Why BNB?

- **A reusable dataset** of published output: not a unique institutional catalogue
- **Uniform format** over 60 years & 3 million records in many languages



The BNB & Linked Data

Selecting resources for linking

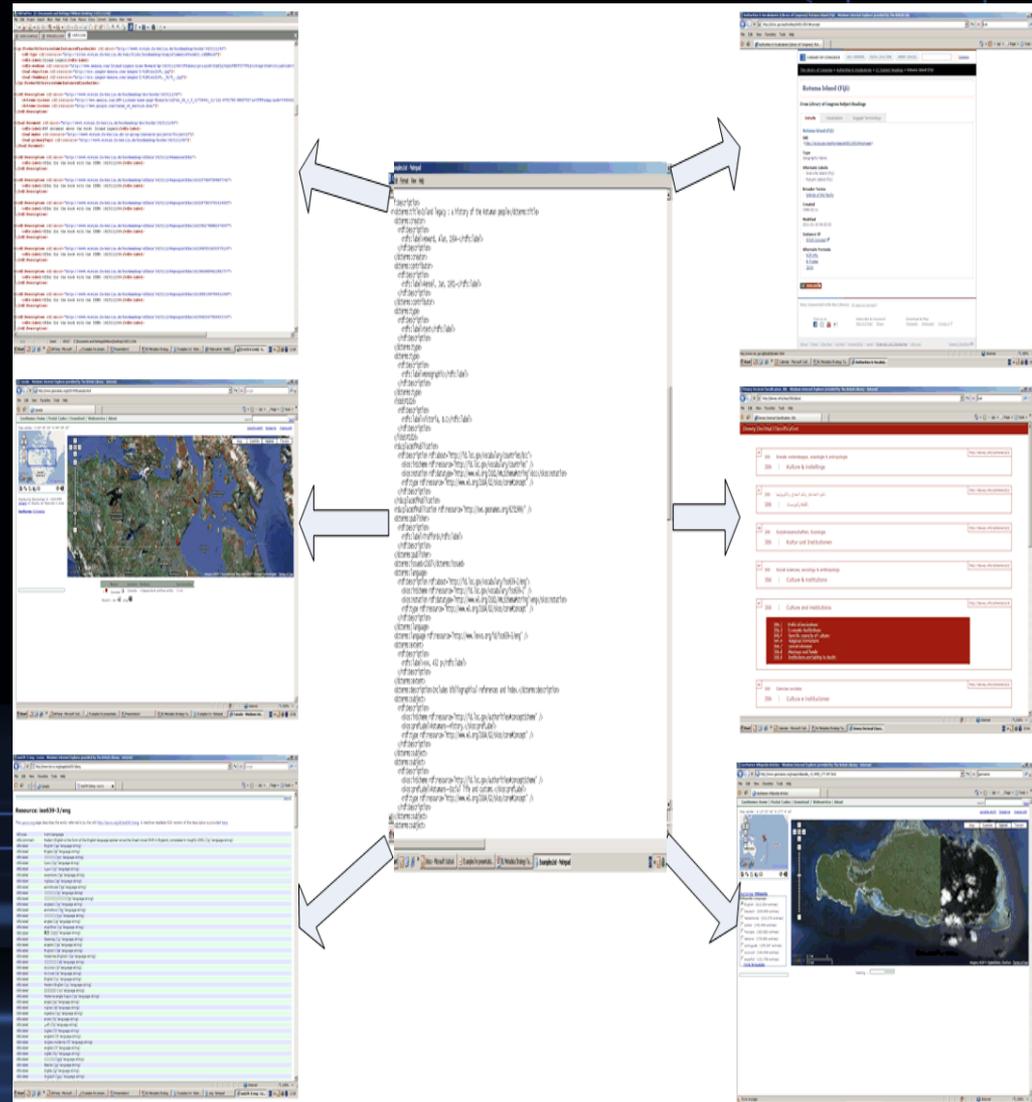
To put BNB data in a wider context

We blended general linked resources:

- *GeoNames*
- *Lexvo*
- *RDF Book Mashup*

With key linked library resources:

- *LCSH*
- *VIAF*
- *Dewey.info*



Collaboration

BL Metadata Services

- BNB Data
- Catalogue Bridge
 - Utilities and tools for manipulating MARC data
- TMQ MARC Global
- Analysts
 - Bibliographic Data
 - MARC
 - FRBR
 - XSLT
- Production team
 - Match & merge
 - Bulk processing

Talis

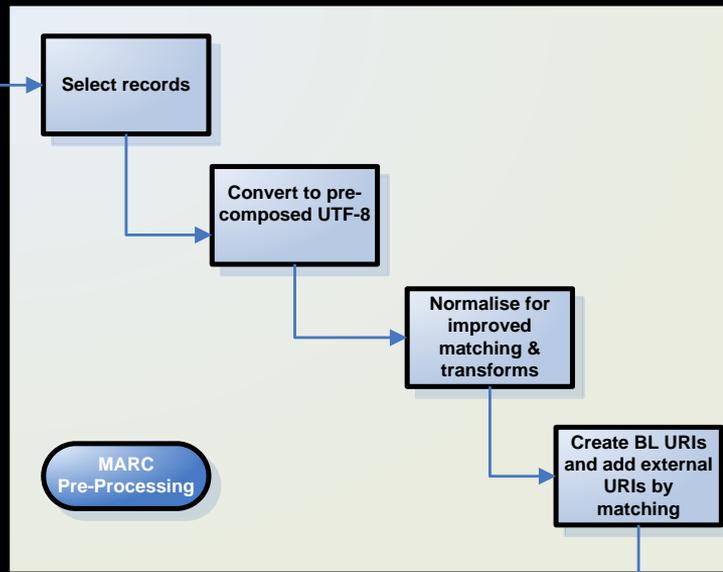
- Training
 - RDF
 - SPARQL
- Technical Infrastructure
- Hosting
- Modelling



The BNB & Linked Data

MARC to RDF Conversion Workflow

Full BNB
MARC21
File



- Selection
- Character set conversion
- Pre-processing
- URI generation

Transform to
RDFXML using
XSLT

BNB RDF/
XML file

Generate RDF
Triple Dump

Load to The Data
Hub (CKAN) etc

Load to Linked
Data Platform

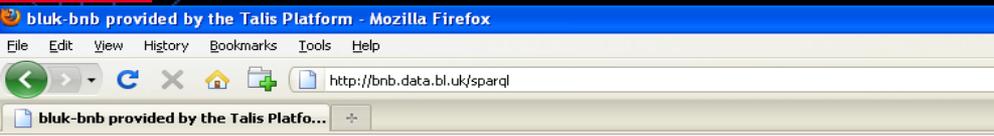
- Data transformation
- Create & load triples

MARC to RDF conversion

Consists of multiple automated steps

Where Did We Get To? Multiple Access Routes

BNB Books 1950-2012
2.6 Million Records
85 Million Unique Triples



bluk-bnb

Run SPARQL query against this store

SPARQL query:



bluk-bnb provided by the Talis Platform - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://bnb.data.bl.uk/search

bluk-bnb provided by the Talis Platform

KASABI BETA

Browse Register Login

Publishing - British National Bibliography (BNB)

BRITISH LIBRARY

British National Bibliography (BNB)
Creative Commons CC0

the Data Hub — The easy way to get, use and share data
Search Groups About

Find datasets

British National Bibliography (BNB) - Linked Open Data

View Resources (21) History

British National Bibliography (BNB) published as Linked Data by the British Library. linked to external sources including VAF, LCSH, Lexico, GeoNames, MARC country, and language, Dewey info, RDF Book Mashup. Published to this data model.

License: Creative Commons CCZero

Tags

- bibliographic
- british-library
- idf
- prosserman
- format-rdf
- format-turtle
- format-json
- library
- iso
- lod
- nonproprietary-vocab
- publications
- published-by-producer

Groups

- BibSoup
- Bibliographic Data
- Library Linked Data
- Linking Open Data Cloud

Resources

- SPARQL Endpoint [api/sparql](#)
- Describe URI Service [api/describe](#)
- Search Service [api/search](#)
- Search Service [api/search](#)
- Example resource - RDF [example/rdf+xml](#)
- Example resource - Turtle [example/rdf+ttl](#)
- Example resource - JSON [example/rdf+json](#)
- Example resource - html [example/html](#)
- Example resource RDF - ttljson/html - default.html [example](#)
- British Library Terms RDF Schema [xml](#)
- Current Information on releases and documentation [html](#)
- BNB_nTriples_jan12_11.zip [application/x.zip-compressed](#)
- BNB_nTriples_jan12_12.zip [application/x.zip-compressed](#)
- BNB_nTriples_jan12_13.zip [application/x.zip-compressed](#)
- BNB_nTriples_jan12_14.zip [application/x.zip-compressed](#)
- BNB_nTriples_jan12_15.zip [application/x.zip-compressed](#)

Published Updated Resources

Bigfoot/4

Search

Search

Maxim

Offset

Sort a

XSL St

Conte

Seat

Bigfoot/4

Explore

View the

Read the

Browse

Talis Platform

- SPARQL endpoint
- Search Service

Kasabi

- SPARQL endpoint
- APIs

The Data Hub

- Triples dump



[Home](#) > [Metadata Services](#) > [Data Services](#) > [Sample Data](#)

Sample Data

Free Data

The following sample files are currently available to download:

Linked Open BNB in RDF/XML

This file provides a sample of linked open data in RDF/XML of published books represented in the British National Bibliography. Data is described using properties from the following vocabularies:

Bibliographic Ontology

Bio: A Vocabulary for Biographical Information

British Library Terms

Dublin Core

Event Ontology

FOAF: Friend of a Friend

ISBD

Org: An Organization Ontology

OWL

SKOS

RDF Schema

WGS84 Geo Positioning

and the preliminary version of **RDA**

[[lodbnbrdfxml](#)] [3.740 KB ZIP Format]

[26 January 2012]

The full triples file is available from [the Data Hub](#).

A copy of the RDF/XML metadata is available from [the Data Hub](#).

Sections:

[British National Bibliography](#)

[Data Services](#)

> [Free data](#)

> [Sample data](#)

[Standards](#)

[News](#)

[Contact us](#)

This page contains links to Adobe PDF files.

Accessibility solutions and free 'Reader' software are available from Adobe.

Beyond the British Library

DC/RDF/XML

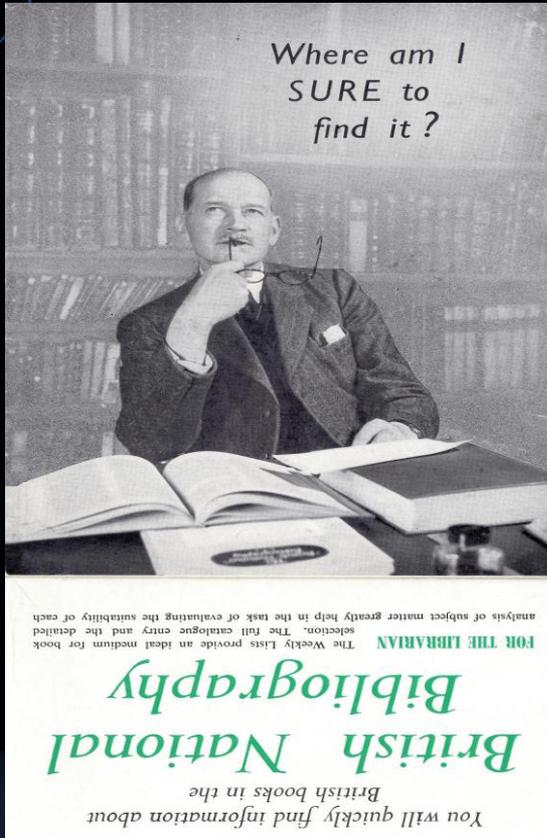
- Open Bibliography Project
- COMET
- Fluidinfo
- ETOC



Linked Open Data

- W3C LLD Incubator Group
- Stanford Linked Data Workshop Technology Plan
 - “...ensure resulting model retains the BL’s high-level focus and its webderived, transparent structure for representing facts about people, organizations, places, events, and topics”
- DBC (Danish Bibliographic Centre)
- ESTC (English Short Title Catalogue)
- sameAs.org

Sustainability: BNB responding to change



Print (weekly lists & cumulations)
MARC Exchange Tapes
CD-ROM

- FTP
- Z39.50
- Online/Web
- PDF (weekly)
- LOD

Sustainability?

Internal resources

Development

Modelling

Data processing

External

Training

Consultancy

Hosting

Kasabi shuts down, says data marketplace 'too slow'

By [Bobbie Johnson](#) | Jul. 9, 2012, 4:48am PT | No Comments

[Tweet](#) 74

[Share](#) 8

[Like](#) 4

[+1](#) 1

[Kasabi](#), a platform that hosted and published linked data, is closing down after

owners Tali growing to

The UK-based to publish d tools to con with each o semantic w moving inf

But after ju opportuniti sustainable. was "time to

Here's the a

“

About

kasabi

Shutting Down Kasabi

Posted on 07/09/2012 by [Leigh Dodds](#)

When we launched Kasabi we set out with a big vision that e and create value from data. Our goal was to create a self-ser publish data using a standard set of APIs, and support deve this we've been moderately successfull: feedback on the site positive. There's a lot of things that I think we got right and Over the last 2 years we've learnt an incredible amount. No about a million and one other things that you encounter wh

However its time to admit that Kasabi is not getting the trac reasons, organisations are not yet ready to fully embrace d in the open data revolution for third-party data marketplac the other startups exploring similar ideas have already beg aspects of their vision.

With the the four and the diff del in in a bet

Sustainability: URIs

- Not principally a technical problem
 - But don't underestimate pain of changing...
- Choices
 - Reuse or roll your own?
 - We generally chose to mint our own rather than rely on external sources
 - sameAs service to manage duplication
 - Transparent or opaque?
 - Transparent was easier for us and developers
 - Opaque more defensible in multilingual context

