Publishing and interlinking music-related data on the web

Yves Raimond, BBC Audio&Music interactive
Linked data
The Web

- Names (URIs)
- Documents (HTML, XML, JSON, ...)
- Interactions with names (HTTP)
Web of documents

- Names identify documents, e.g. HTML, XML, etc.
- Documents are interlinked:

  <a href="http://moustaki.org/" />
Web of documents
We understand...
We understand...

## Top Artists

<table>
<thead>
<tr>
<th>Position</th>
<th>Artist</th>
<th>Listen Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Fall</td>
<td>244</td>
</tr>
<tr>
<td>2</td>
<td>PJ Harvey</td>
<td>157</td>
</tr>
<tr>
<td>3</td>
<td>Magoo</td>
<td>131</td>
</tr>
<tr>
<td>4</td>
<td>Roots Manuva</td>
<td>127</td>
</tr>
<tr>
<td>5</td>
<td>Sean Paul</td>
<td>103</td>
</tr>
<tr>
<td>6</td>
<td>Quasimoto</td>
<td>101</td>
</tr>
<tr>
<td>7</td>
<td>Joy Division</td>
<td>97</td>
</tr>
<tr>
<td>8</td>
<td>Sizzla</td>
<td>97</td>
</tr>
<tr>
<td>9</td>
<td>Radiohead</td>
<td>82</td>
</tr>
<tr>
<td>10</td>
<td>Jawbone</td>
<td>78</td>
</tr>
<tr>
<td>11</td>
<td>Bonzo Dog Band, The</td>
<td>75</td>
</tr>
<tr>
<td>12</td>
<td>New Fast Automatic Daffodils</td>
<td>72</td>
</tr>
<tr>
<td>13</td>
<td>Rosinators, The</td>
<td>70</td>
</tr>
<tr>
<td>14</td>
<td>Yellowman</td>
<td>69</td>
</tr>
<tr>
<td>15</td>
<td>Destiny’s Child</td>
<td>66</td>
</tr>
</tbody>
</table>
Machines don't...
Shortcomings

- Opacity of documents
  - “In 1883, Fauré married Marie Fremiet”

- Untyped links
  - “Edgar Faure was born in Béziers”
  - Interpreting the surrounding context is hard for a machine

We're unable to use the Web as a big database
Silos and views
Linked data

Let's expose and interlink the actual data, not only the documents!
Linked data

The 4 Linked Data principles (TimBL 06)

1) Use URIs as names for things
2) Use HTTP URIs so that people can look up those names
3) When someone looks up a URI, provide useful structured data
4) Include links to other URIs, so that they can discover more things
RDF

- RDF is the web standard for such structured data

- RDF data model:
  - Subject (URI)
  - Property (URI)
  - Object (URI or literal)

- “Fauré” “born in” “London”
- “I” “work near” “Broadcasting House”
And that's all! (yes, really...)
Web of data

A web of things
Linking Open Data project

- Use these technologies to create a “web of data”
- Publish and interlink various open datasets, across various domain and disciplines
  - Music,
  - Bio-informatics,
  - Scientific publications,
  - Libraries,
  - Geography
  - Census data...
How are they related?

(No, not the moustache)
Gabriel Fauré

http://dbpedia.org/resource/Gabriel_Fauré
Is the same as

http://rdf.freebase.com/ns/guid.9202a8c04000641
Was born in

http://rdf.freebase.com/ns/en.pamiers
Is the same as

http://dbpedia.org/resource/Pamiers
Is located within

http://dbpedia.org/resource/Ariège
Is located within

http://dbpedia.org/resource/Midi-Pyrénées
Is the birth place of

http://dbpedia.org/resource/Charles_Boyer
Is part of

http://dbpedia.org/resource/Category:Drug-related_deaths_in_Arizona
Has member

http://dbpedia.org/resource/Mattie_Blaylock
Is spouse of

http://dbpedia.org/resource/Wyatt_Earp
Why is that important?

- Data without context is almost useless
- Make your data accessible for other researchers
- Make your data linkable
- Publish and maintain only your data, and leave the burden of ancillary descriptions to other communities of interest
- … who link to other communities of interest …
- Enable cross-domain queries to be formulated
- “The whole is greater than the sum of the parts”
The Music Ontology
The Music Ontology

- A framework to deal with music-related information on the web of data
- Based on other ontologies
  - Timeline
  - Event
  - FOAF
  - FRBR
The Timeline ontology

- Expressing temporal information on multiple timelines
  - This performance happened on the 31\textsuperscript{st} of August, 2008
  - There is a beat occurring around sample 32480
  - The second verse is just before the second chorus

- Timelines for e.g.
  - Scores
  - Performances
  - Audio signals
The Timeline ontology

a) instant

timeline

audio signal timeline

start

PT3S

b) interval

timeline

universal timeline

start

2001-10-26T12:00:00Z

duration

P7DT
The Event ontology

- Classifying space–time regions
  - This performance involved Glenn Gould playing the Piano
  - This signal was recorded using a Shure microphone located on the top–left of the guitar amplifier
  - On this particular recording of that performance, that particular part starts at 12 minutes and 53 seconds and lasts 9 minutes 12
The Event ontology

- geo:SpatialThing
- time:TemporalEntity
- foaf:Agent
- Thing
- Event

Relationships:
- sub_event
- place
- time
- factor
- agent
- product
- Thing
FRBR

- Functional Requirements for Bibliographic records
- Layering in abstraction
  - Work
  - Expression
  - Manifestation
  - Item
The Music Ontology...

...subsumes these ontologies to deal with music-related information
The Music Ontology

- Editorial
  - Record, Track, Playlist, Vinyl, CD, Artist, Band
- Production workflow
  - Musical work, Arrangement, Score, Performance, Sound, Recording, Audio Signal
- Roles
  - Engineer, Performer, Composer, Listener
- Anchor points for taxonomies
  - Instrument, Genre, Form
The Music Ontology
But also...

- Extensions to deal with:
  - Chords and keys
  - Audio features, e.g. Key changes
  - Instrument taxonomies
  - Symbolic information
  - Visualisation tools
    - [http://www.sonicvisualiser.org](http://www.sonicvisualiser.org)
  - LOTS of available linked data!!
    - [http://dbtune.org/](http://dbtune.org/) (>14 billion statements)
    - [http://www.bbc.co.uk/music](http://www.bbc.co.uk/music) (~500 million)
Linked Data at the BBC
Not so long ago...

- Static web pages uploaded via FTP
- Strictly hierarchical navigation
Things are changing...

- It's all about the resource and its relationship to other things, within the BBC or outside
  - Artists
  - Programmes
  - Events
  - Places
  - Reviews
  - Food
  - News
  - ...
What we're doing

- One URI for everything
- Different representations for a single thing
  - HTML
  - RDF
  - JSON
  - Mobile
  - Etc.
- Links to other things within these representations
- From a tree to a graph...
Ontologies at the BBC

- [http://www.bbc.co.uk/ontologies](http://www.bbc.co.uk/ontologies)
- Define the things we publish a page for
  - Programmes
  - Categories
  - Artists
  - Etc.
- Define the links between those things
  - Programme → Artist
  - Artist → Programme
  - Programme → Category...
Programmes BETA

Welcome
This site aims to ensure that every programme the BBC broadcasts has a permanent, findable web presence.

Find out more about this site, discover why this is a BETA and take a peek at our plans for the future.

Programmes for you
Whether it's games & quiz shows, all new drama or the full suite of music, we've got the programmes for you.

Explore the links below to find programmes you're interested in. Browse by name, category or schedule.

A to Z
Use the A to Z to find a programme by name...

Categories
We categorise our programmes by genres and formats. Browse these categories and discover recent programmes of interest...

Schedules
We're adding the ability to browse via schedules. As these become available we'll link to them below.
Demo

Re-purposing our data:

http://api.talis.com/stores/bbc-backstage/services/sparql

http://blog.dbtune.org/

http://api.talis.com/stores/ldodds-dev1/items/subject-relations.html

http://api.talis.com/stores/ldodds-dev1/items/reviewers.html
Questions?

SEMANTICS

TO THE PEOPLE