Towards a musical Semantic Web

Yves Raimond
Centre for Digital Music, Queen Mary, University of London

May 6th, 2007

Queen Mary
University of London

centre for digital music

Overview

Queen Mary University of London

- **Introduction**
  - Web
  - Web of data
- **The Semantic Web**
  - RDF
  - Making sense of the data
  - Content negotiation
- **The Music Ontology**
  - The Timeline ontology
  - The Event ontology
  - FRBR + FOAF
  - Music *production* specific concepts
  - Workflow information
  - Levels of expressiveness
  - Extensions
- **And now??**
  - Linking open data on the Semantic Web
  - Two applications

Introduction – Web

Looking for Creative Commons-licensed song from the French band *Lonah*

1. I ask my favourite search engine for "*Lonah creative commons song*"

Tryad – Music at Last.fm
Tagged as: **creative commons**, electronic, ambient. People who like Tryad also like Ehma, Revolution Void, **Lonah**. Learn more about Tryad at Last.fm, **...**
www.last.fm/music/Tryad/+journal - 108k - Cached - Similar pages

Lonah – Music at Last.fm
Learn more about **Lonah** at Last.fm, the world's largest social music platform. **...** Play piano Tag Radio Play in pop up · Play **creative commons** Tag Radio Play **...**
www.last.fm/music/**Lonah** - 85k - Cached - Similar pages
[More results from www.last.fm]

annma's blog
jamendo.com is also a very good place to download **creative commons** music, **...** p.o. box, try^d, drunksouls, thierry blanchard, -=kwada=-, **lonah**, and more :) **...**
annma.blogspot.com/2006/06/free-music.html - 11k - Cached - Similar pages

Internet Archive: Details: Black Sweater White Cat 11.26.2005
Program in focused on copyleft and **creative commons** music by the people, **...** http://www.wmrecordings.com/ **song**: Artiste album: PiÃ¨ces artist: **Lonah ...**
www.archive.org/details/bswc11262005 - 21k - Cached - Similar pages

Internet Archive: Details: Black Sweater, White Cat 01.28.06
Black Sweater, White Cat is focused on copyleft or **creative commons** licensed music from around the internet **...** This edition is 100% **Creative Commons** music. **...**
www.archive.org/details/BSWC012806 - 20k - Cached - Similar pages

Introduction – Web

Looking for Creative Commons-licensed song from the French band *Lonah*

2. I read the *context* of each of the first results
3. The second one seems ok...
4. I reach this *last.fm* page:

Lonah

20,015 plays scrobbled on Last.fm

Lonah is a young French rock band. Music available on http://www.jamendo.com/us

Edit this artist description

User Tags (see more)

cabaret creative commons dreamy female vocalists **french** jamendo piano rock

Tag this artist

5. According to the tags, it looks like the band I am looking for...
6. I read "Music available on ..." and decide to visit the linked page
7. I reach the Jamendo website
8. I launch a search for *Lonah*, and, finally:

	5	Crepuscule	4:20
	6	Les amants de cristal	5:54
	7	Les effacés	4:05
	8	Fractale	3:47

Introduction – Web

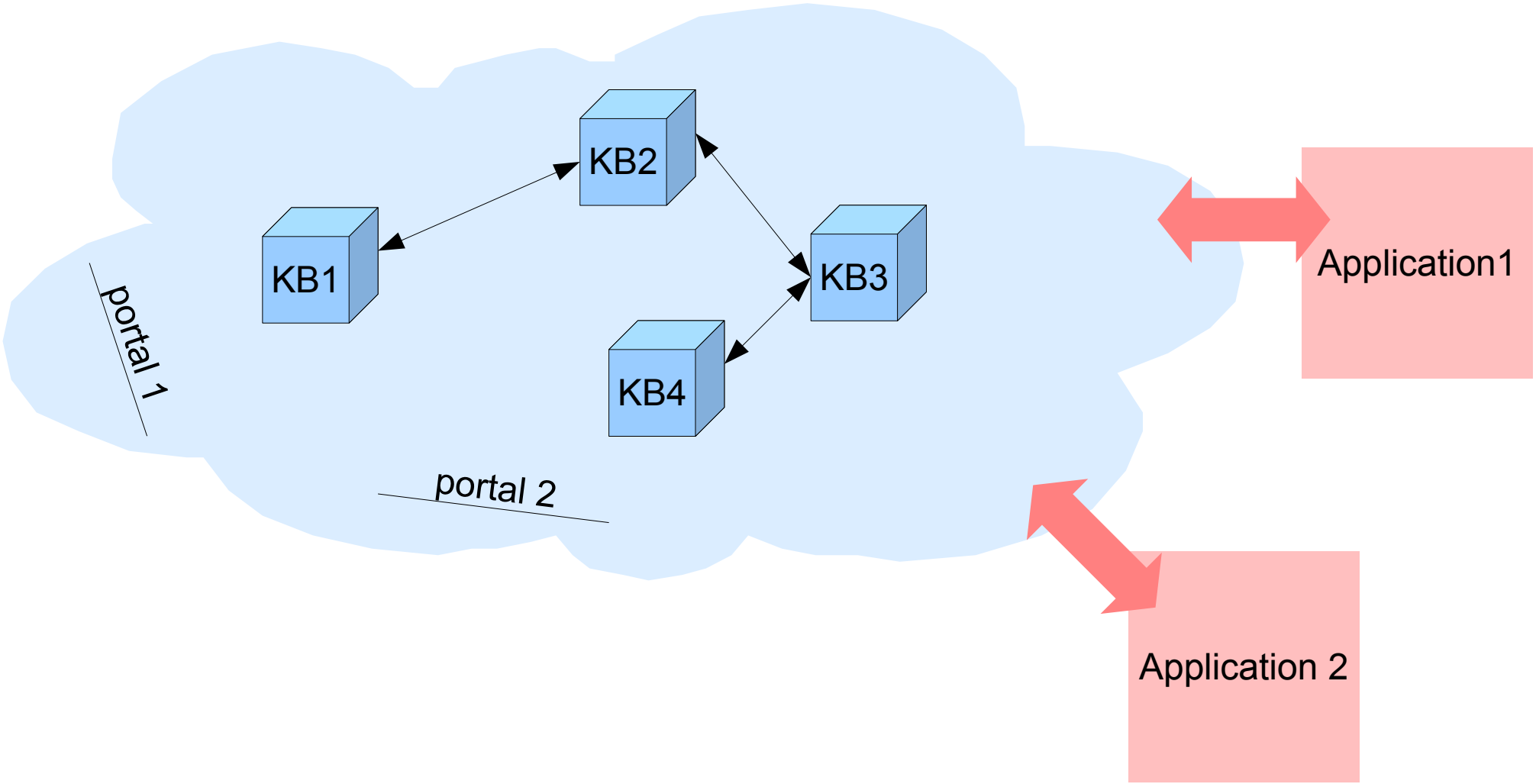
Some requirements emerging from this scenario:

- I need an entry point: the **search engine**
- I need to understand the **context** of the links
- I need to find my way into the **web maze**

Now: **Ask your computer to do the same thing!**

Introduction – Web of data

Turning the Web into a huge, “semantic”, democratic database in order to make machines able to look by themselves for particular informations



The Semantic Web

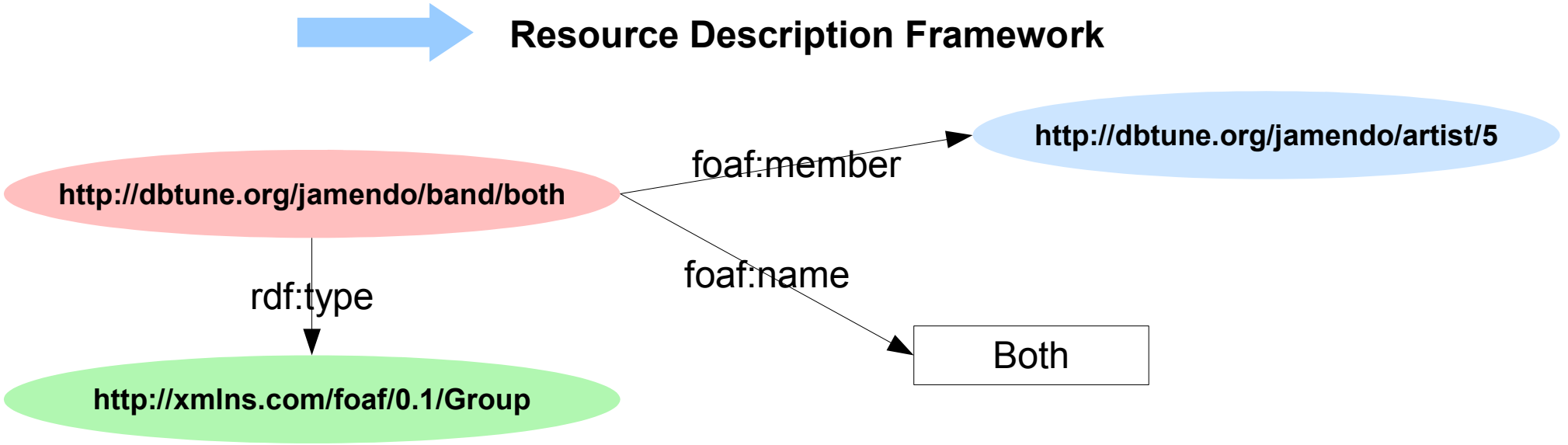
Queen Mary University of London

Resources on the Web can be far more than just web pages!

http://moustaki.org/foaf.rdf#moustaki is a resource representing *me*

http://dbtune.org/jamendo/band/lonah is a resource representing the band *Lonah*

When HTTP-GETting, Let's leave fancy HTML pages for humans, and let's provide some useful descriptions for the machine!



Ontologies - Making sense of the data

Ontologies, to map these **resources and properties (links)** to **real-world objects and relationships**

Providing a COMMON UNDERSTANDING

An **Album** has several **Tracks**, a **name**, a **release date**...
A **Performance** has one **location**, one **time**, some **performers**, ...

Ontologies are also described in **RDF**

Instance data **refers** to ontologies through RDF triples such as:

- <http://dbtune.org/jamendo/artist/5> rdf:type <http://purl.org/ontology/mo/Musicartist>
- <http://dbtune.org/jamendo/artist/5> foaf:name “Both”

Content negotiation

And now, let's make both the **human** and the **machine** happy!

http://dbtune.org/jamendo/artist/5

HTML for “human consumption”

```
<mo:MusicArtist
rdf:about="http://dbtune.org/jamendo/artist/5">
   <foaf:based_near
rdf:resource="http://dbpedia.org/France"/>
  <foaf:homepage rdf:resource="http://www.both-
world.com"/>
  <foaf:img
rdf:resource="http://img.jamendo.com/artists/b/both.j
pg"/>
  <foaf:name
rdf:datatype="&xsd;string">Both</foaf:name>
</mo:MusicArtist>
```

RDF for “machine consumption”

The Music Ontology

Problem: no agreed ways of dealing with music-related information on the Semantic Web

Solution: Let's launch a community project, based on previous ontology engineering efforts!

http://musicontology.com/

Several facets [Pachet]:

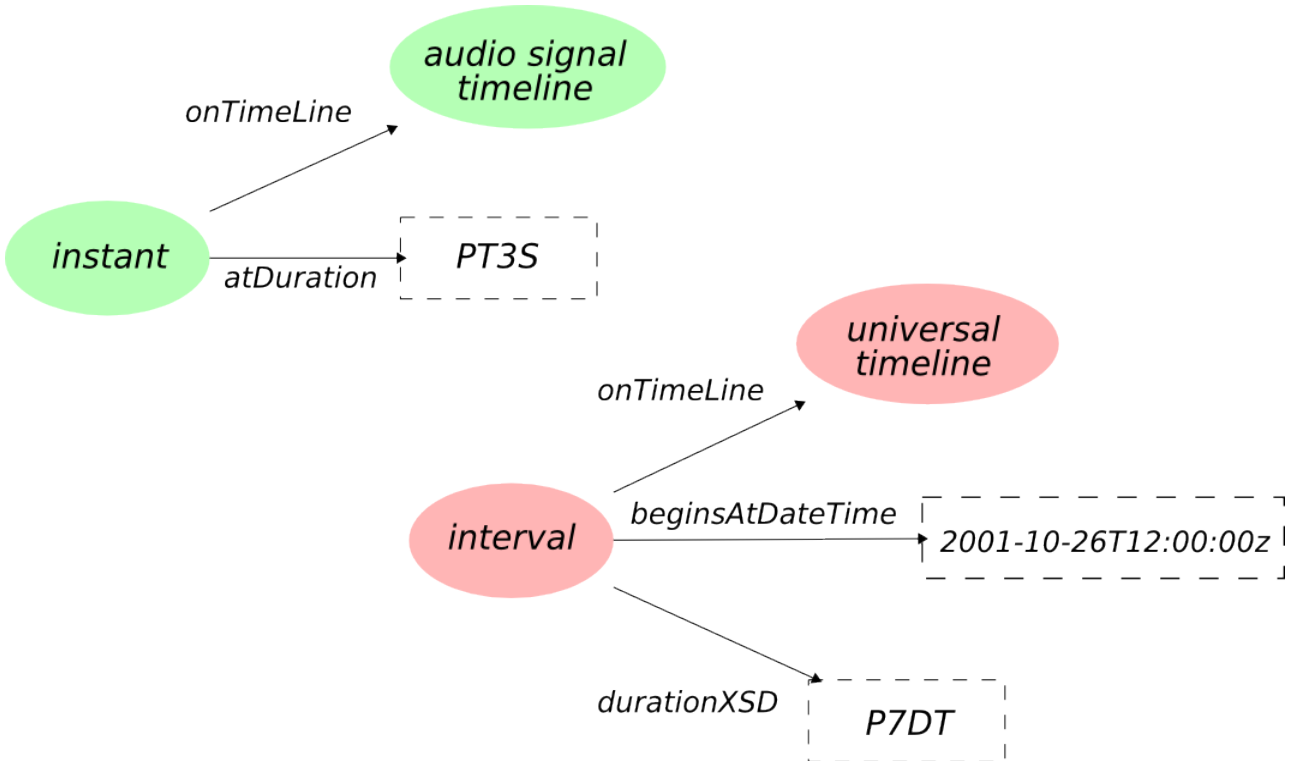
- Complex *editorial* information
- *Acoustic* information
- (*cultural* information)

The Timeline ontology

First thing to address: representing *temporal information*

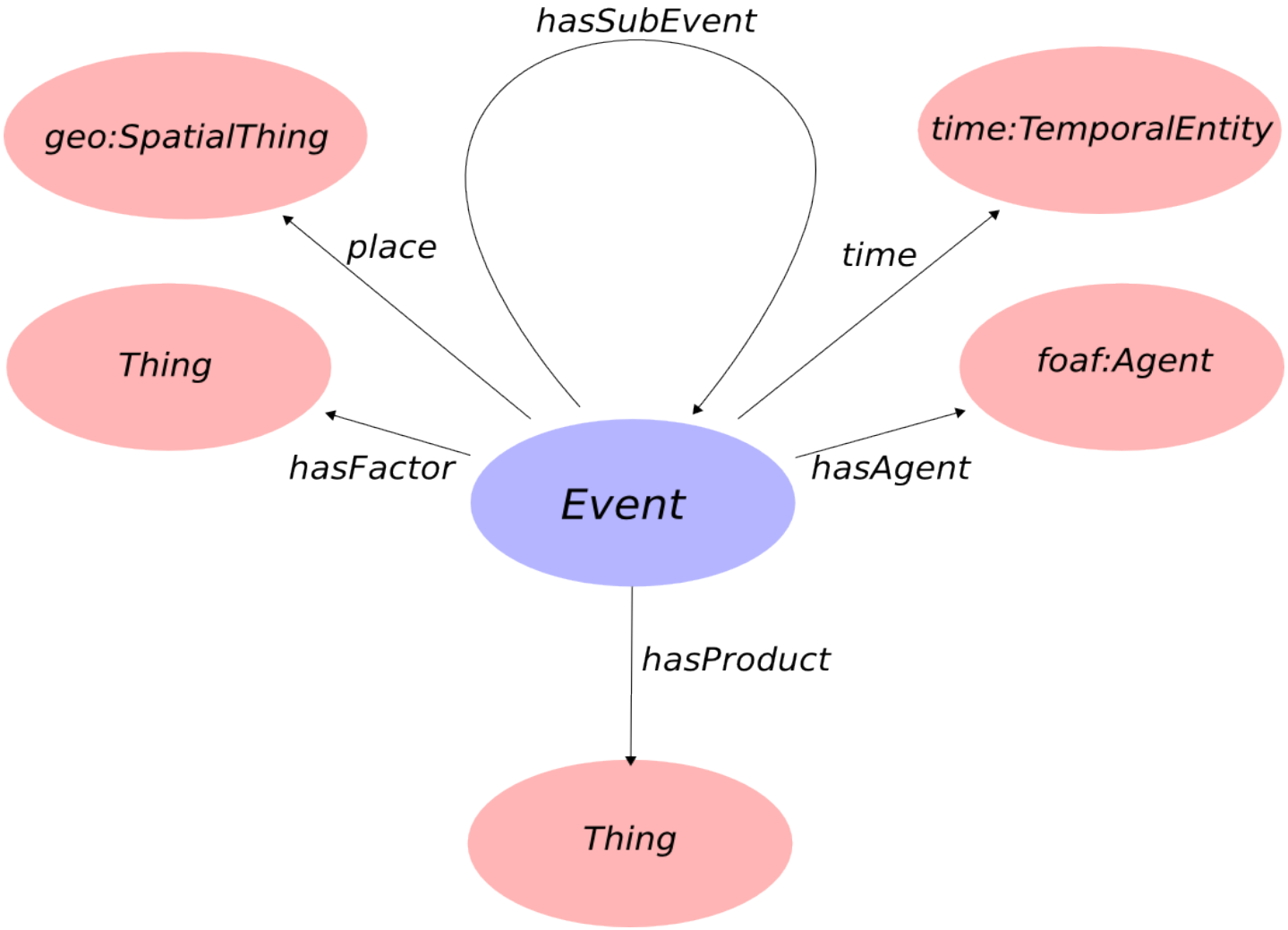
"This performance happened the 9th of March, 1984"
"This beat is occurring around sample 32480"
"The second verse is just before the second chorus"
...

Only four concepts: **Instant**, **Interval**, **TimeLine** (and **TimeLineMap**)



The Event ontology

We need a way to classify space/time regions :
Performances, recordings, beats, verses, composition, ...



FRBR + FOAF

Queen Mary University of London

FRBR: Functional Requirements for Bibliographic Records

We use three FRBR concepts:

- **Work**
- **Manifestation**
- **Item**

The **Expression** concept seemed to fuzzy for being used:
whole *workflow* between a work and its manifestation

FOAF: Friend-of-a-friend

- **Person**
- **Group**
- **Organization**

- ... and the relationship vocabulary (married, brother of, etc.)

music ontology

Music *production* specific concepts

On top of FRBR:

MusicalWork, **MusicalManifestation** (**Album**, **Track**, **Playlist,** etc.)
MusicalItem (**Stream**, a particular **Vynil**, etc.)

On top of FOAF:

MusicArtist and **MusicGroup** (defined classes)
Arranger, **Engineer**, **Performer**, **Composer**, etc. (same thing)

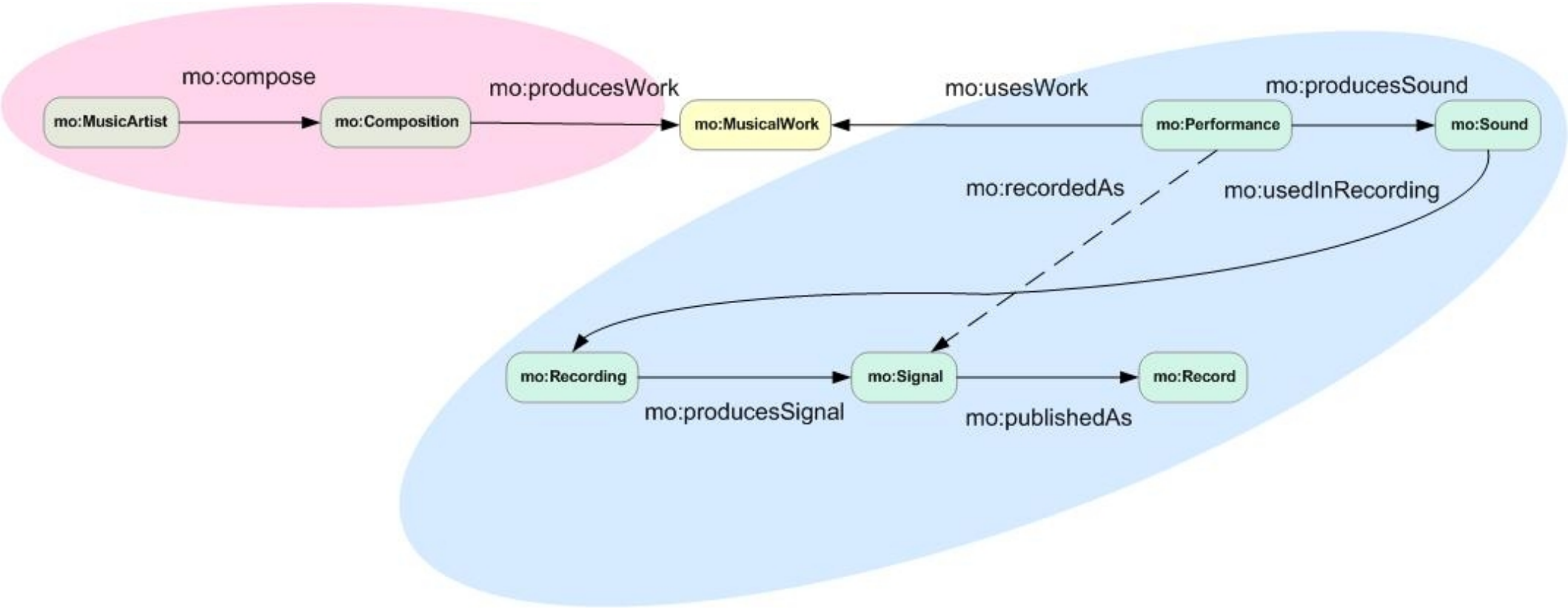
On top of the Event ontology:

Composition, **Arrangement**, **Performance**, **Sound**, **Recording**

Others:

Signal, **Score**, **Genre**, **Instrument**, etc.

Workflow information



Levels of expressiveness

Flexibility of the ontology

- Level 1: **purely editorial**

  "This track is on that particular album and that compilation and was created by that artist"

- Level 2: **introducing events**

  "This is a recording of this particular musician playing that jazz-rock arrangement of that particular piece"

- Level 3: **introducing event decomposition**

  "In this performance, this key was played at this particular time by this person, who was playing the piano"

Extensions

Lots of **anchor points** (instrument, genre, signal, timeline, etc.)

Already several extensions available:

- **Musical feature ontology**: uses **Event** as a way to classify features on a signal' timeline
- **Instrument taxonomy**: thanks to Musicbrainz!
- **Genre taxonomy**: thanks to Wikipedia/DBPedia!
- **The Key ontology**

Other possible extensions:

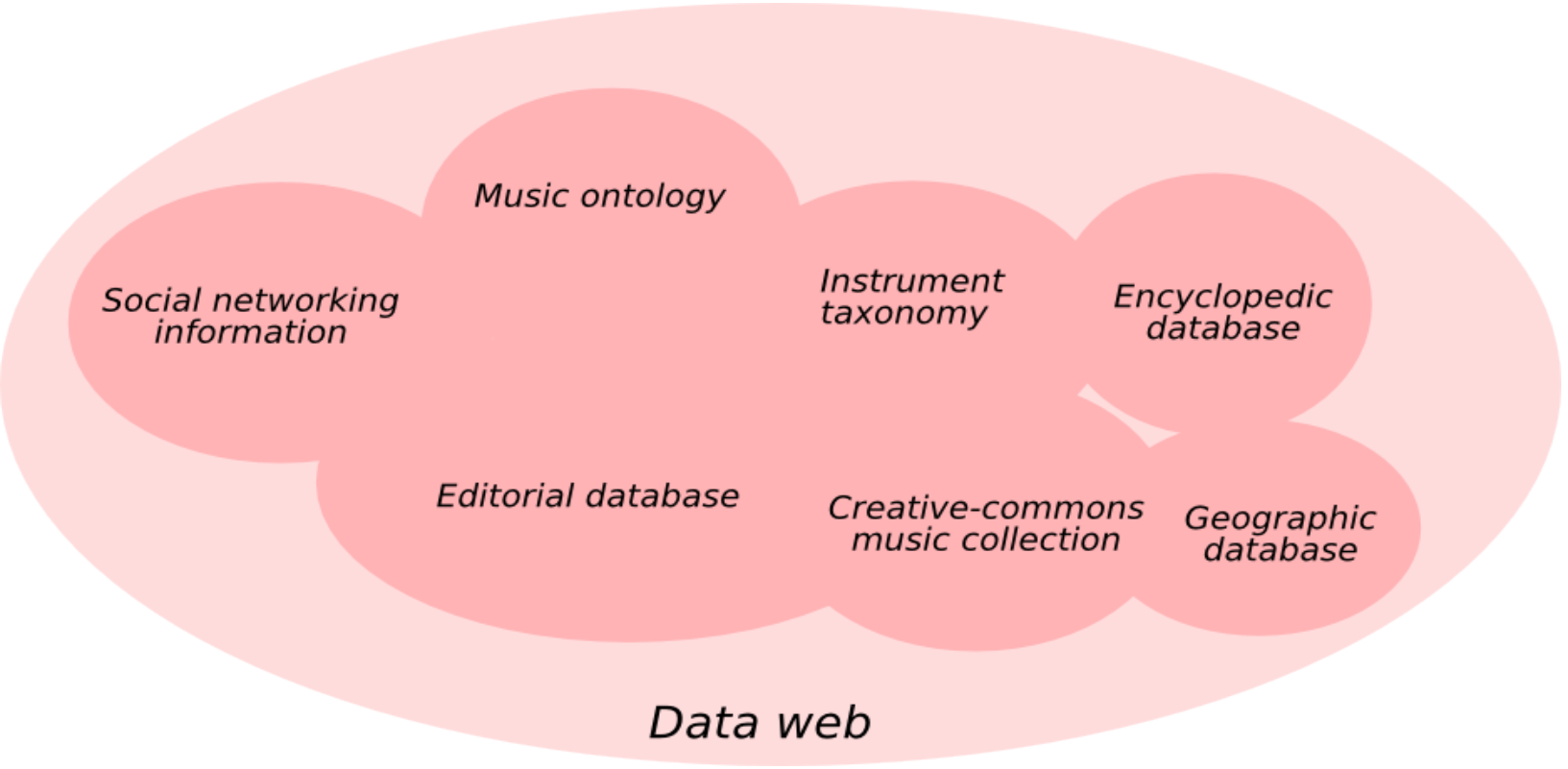
- Audio recording devices under the **Recording** concept?
- **Mixing** events dealing with **Signal** objects?
- Sound cognition under the **Sound**/**Listener** concepts?
- Symbolic music notation under **Score**?
- Chord ontology?

Linking open data on the Semantic Web

W3C' Semantic Web Education and Outreach community project

Lots of *open data* available: Wikipedia, Geonames, Musicbrainz, creative commons repositories, etc.

Let's interlink them using Semantic Web technologies: DATA MASHUPS



So far:
- Jamendo
- Magnatune
- Musicbrainz
- DBPedia
- GeoNames
- RDF book mashup
- ...

And now??

Management of audio collections and enhanced access:

- Your audio files are just other **items** of a particular **manifestation**, which has an URI
- Store the corresponding statements in your SW-enable application
- And your collection gets access to the whole web of knowledge (well, in its current state:-))

Give me all musical works composed in a city with more than 500 000 inhabitants

Is there someone nearby really liking this band and the same beer as me, so that we can have a drink tomorrow?

Place my collection on a timeline and make me listen something composed in the UK in 1560, followed by a rock song recorded in the 60s

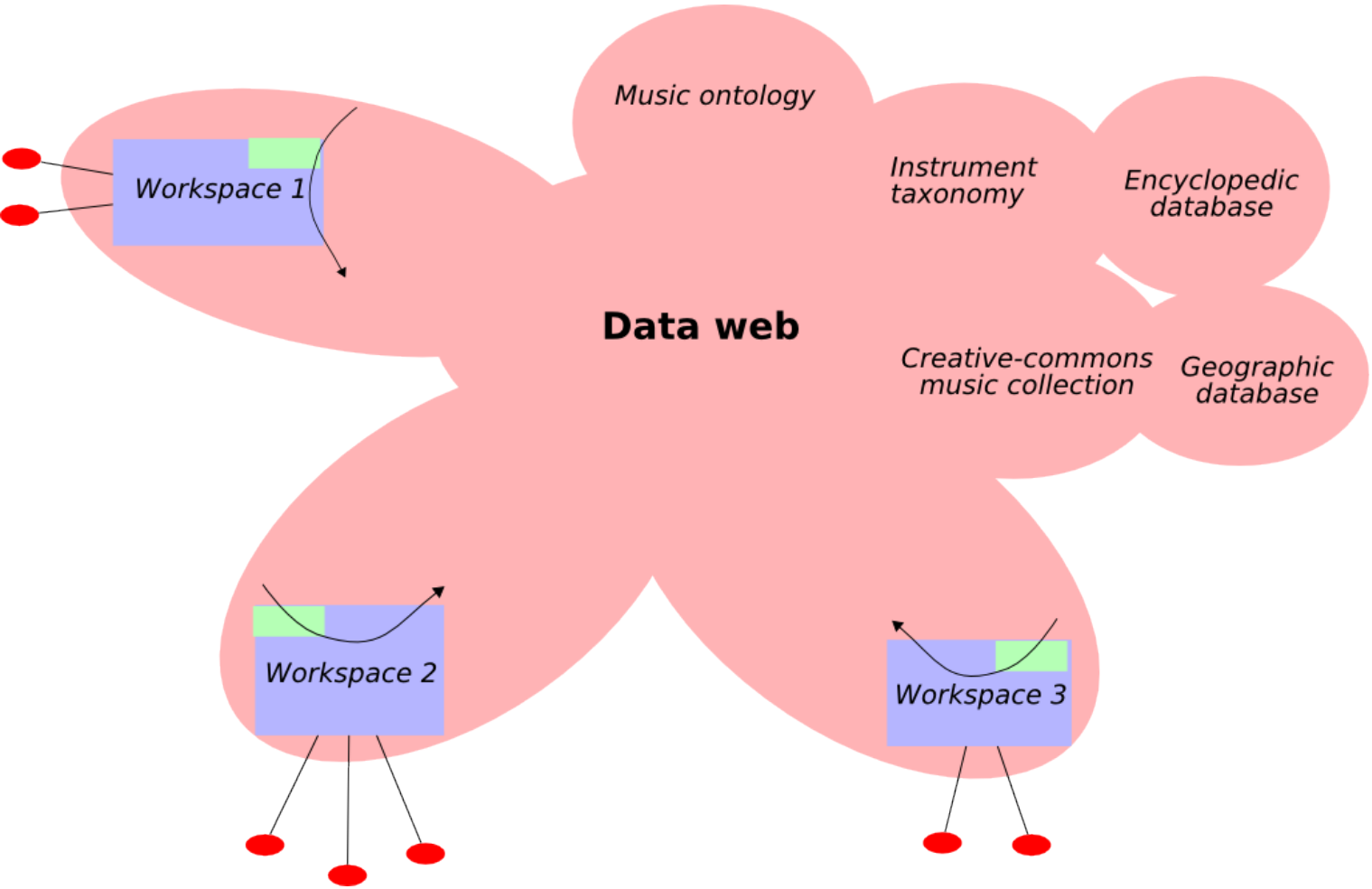
Give me all Jimmy Hendrix songs played by Brass Bands with at least 5 members

Are there any other performances of this work? Give me one with a small part at 120 bpm

***Semantic workspaces*:**

- This web of data is indeed a **machine-understandable cultural web**
- Let's use this artificial culture to make algorithms **smarter than they are**, and export resulting knowledge dynamically!

Semantic workspaces



: external data processing tools
: `semantic workspace'
: semantic-web client
: data web
: knowledge processing workflow

Thank you!!