1 Ontology matching

So far, I have looked for artist matches between these ontologies:

1. LODAC
2. DBPedia
3. Japanese Wikipedia Ontology (Keio university)

Using simple string matching, I was able to map this number of artists between each:

- LODAC → Japanese DBPedia: 130 matches
- LODAC → English DBPedia: 279 matches
- LODAC → Keio Ontology: 790 matches

However, even though the Keio Ontology is based on Japanese Wikipedia and has 11685 artists, I could only find 722 matches with Japanese DBPedia (870 for English DBPedia). There must be a way to find more.

2 How useful is the Keio taxonomy?

The Keio ontology does not break the artists up well. Most of the artists are within just three categories. This means browsing by category based on this hierarchy is hard.

Also, it is focused too much on Japanese artists. I would like to make a hierarchy that makes all classifications of artists browsable, so that the links between them can be shown.

3 Automatically creating a taxonomy of artists

I am trying to use **hierarchical clustering** (complete linkage clustering).

This looks at all of the links between subjects to determine the strongest:

$\max \{d(a, b) : a \in A, b \in B\}$

$d(a, b)$ (distance between each category) is determined by:

For each pair of categories:

- Does an artist belong to both categories? (strong link)
- If two artists are in category A, does one belong to category B? (weak link)
- Do both categories appear in Japanese DBPedia's WikiLink field for any artists (strong link)