

## Practical session: Data linking with LINES

This practical session is dedicated to the use of a well-known tool that is LINES.

### 1) Hands on LINES

Go to the lab link at e-campus (<https://ecampus.paris-saclay.fr/course/view.php?id=38304#section-4>)

1- Download and unzip the archive [LINES.0.6.RC4-2.zip](#)

2- Open a terminal and go to the folder [LINES.0.6.RC4-2](#)

3- Run the command line:

```
java -jar LINES.jar ./Release_Examples/dbpedia-actors.xml
```

What is the runtime? .....

4- Open in a text editor the file `./Release_Examples/ dbpedia_actors_reviewme.nt`

How many lines you get? .....

5- Open in a text editor the file `./Release_Examples/dbpedia-actors.xml`

What is the similarity measure that is used?

6- Modify this similarity measure with [trigrams](#) measure and run again the command line:

```
java -jar LINES.jar ./Release_Examples/dbpedia-actors.xml
```

What is the runtime? .....

7- Open in a text editor the file `./Release_Examples/ dbpedia_actors_reviewme.nt`

How many lines you get? .....

## 2) Linking conference data (to deliver via e-campus)

Download cora dataset from (big file):

[https://www.lri.fr/~sais/dev/rdf/SplitedDataModif3/concept-graph-original.rdf\\_modif.rdf](https://www.lri.fr/~sais/dev/rdf/SplitedDataModif3/concept-graph-original.rdf_modif.rdf)

A **small** part of it can be Downloaded from:

[https://www.lri.fr/~sais/dev/rdf/SplitedDataModif3/concept-graph-original\\_sp1.rdf\\_modif.rdf](https://www.lri.fr/~sais/dev/rdf/SplitedDataModif3/concept-graph-original_sp1.rdf_modif.rdf)

- 1- Write a LIMES specification file by taking inspiration from the examples above and the user-manual that you can find at: <https://ecampus.paris-saclay.fr/mod/resource/view.php?id=353043&redirect=1>
  - Use several properties to compare conferences
  - Use different similarity measures
- 2- Run LIMES on the small dataset of cora (you can use the same dataset as a source and as target just give them different names).

What are the results?

- 3- Try to change some parameters in the specification file (similarity measures, the linking thresholds)

What are the results?

- 4- Run LIMES on the big dataset of cora  
What are the results?

**Deliver on the e-campus course all the <XXX.nt> files that you obtain.**