

## Manual for the Cortext practical session

March 12<sup>th</sup> 2018

Antoine Schoen

- I) For uploading the database « OECD\_TIP\_prepared\_corpus\_final.txt.zip »



Click on “upload a new corpus”

And select the following options:

- Type of data: dataset
- Corpus format: robust csv
- Add: YEAR as time entry (in capital letters)

SCRIPT PARAMETERS

Source

Type of Data

☒ dataset
 ☐ term list
 ☐ cortext db

Corpus Format

Please indicate the format of your csv file

☒ tabulation
 ☐ colon
 ☐ semi-colon
 ☐ pipe

If certain columns have multiple values, please indicate the intra-field separator

If certain columns have multiple embedded values, please indicate the secondary intra-field separator

separator

If your csv file includes a time entry, please indicate the attribute name

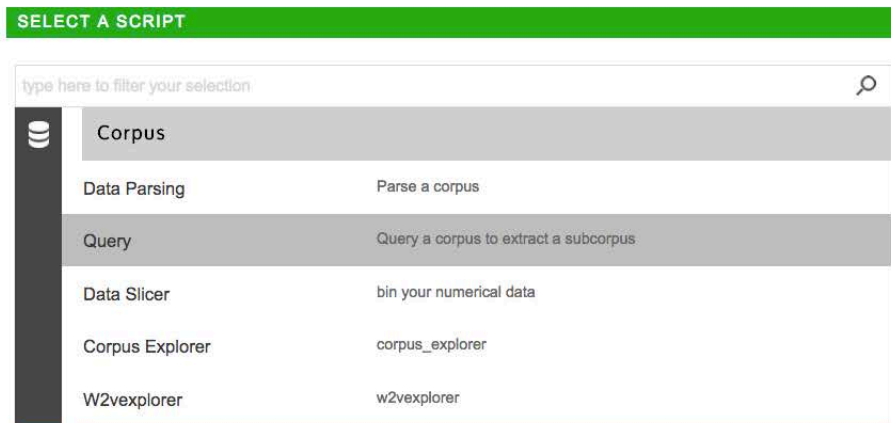
By default, time entry should be formatted as an integer, activate this option if you have dates

II) For reducing the DB to the 1994-2014 period

Click on “Start a new script”



In the Corpus section, select “Query”



Restrict to the 1994-2014 period

**Query Parameters**

**Query Type**

☒ sql ☐ pivot field

**target table name** YEAR

**condition** data <> "1993" AND data <> "2015"

**Inverse the condition**



☐ yes ☒ no

**Build an entirely new database excluding articles that do not match the query for every fields (\*)**

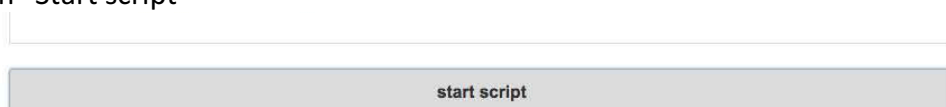
☒ yes ☐ no

**expert sql query**

**output table/bdd name** corpus\_1994\_2014

© INRA 2016 | Contact | Crédits | Mentions Légales | CGU  

Click on “Start script”

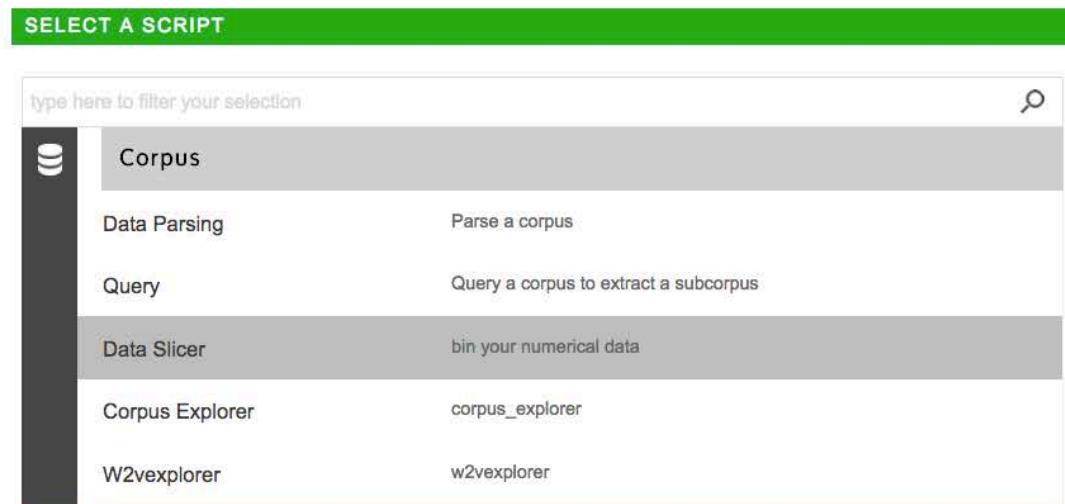


III) For defining 3 time slices: (1994-2000), (2001-2008), (2009-2004)

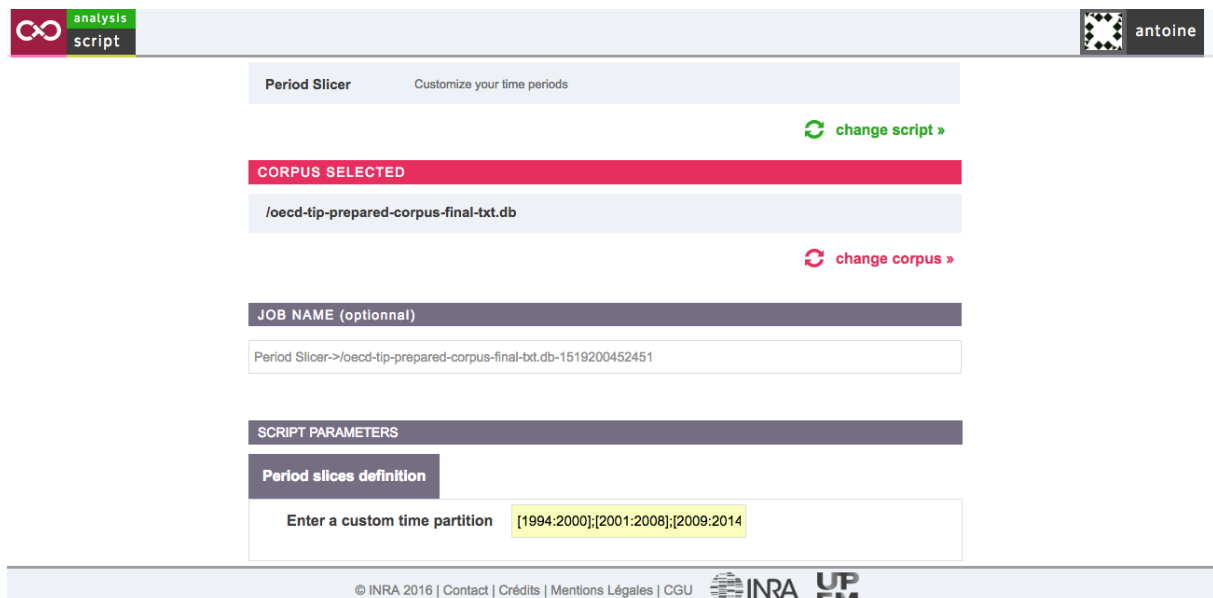
Click on “Start a new script”



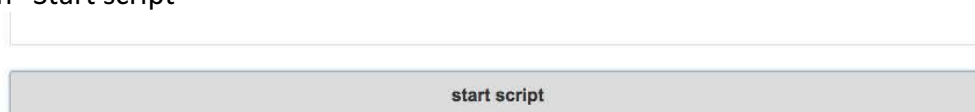
In the Corpus section, select “Period Slicer”



Define the 3 time slices (1994-2000), (2001-2008), (2009-2004)



Click on “Start script”

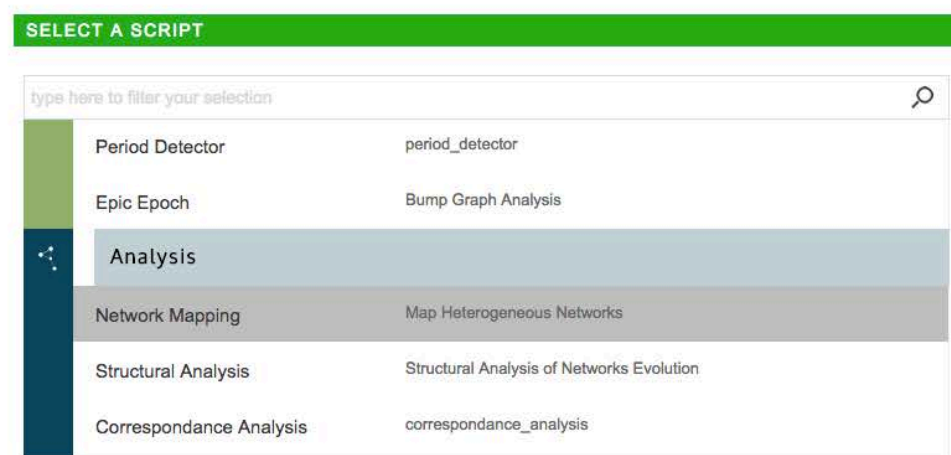


IV) For producing the maps

Click on “Start a new script”



In the Analysis section, select “Network mapping”



Choose the following script parameters

- In **Nodes Selection**, select: First field CLEANED\_TERMS; Second field CLEANED\_TERMS; Number of Nodes: 150

SCRIPT PARAMETERS	
<b>Nodes Selection</b>	Edges   Dynamics   Network Analysis and layout
First Field	CLEANED_TERMS
Second Field	CLEANED_TERMS
Number of nodes	150
Nodes advanced settings	no

- In **Dynamics**, choose: Custom Period; Number of time slices: 3; Time slices distribution: regular

SCRIPT PARAMETERS

Nodes Selection

Edges

Dynamics

Network Analysis and layout

Choose Original Timescale

☐ Standard Periods

☒ Custom Periods

Number of time slices

3

time slices distribution

☒ regular

☐ homogeneous

☐ smoothed

The dataset will be divided in n periods

Overlapping periods

☐ yes

☒ no

sequencing

☒ snapshot

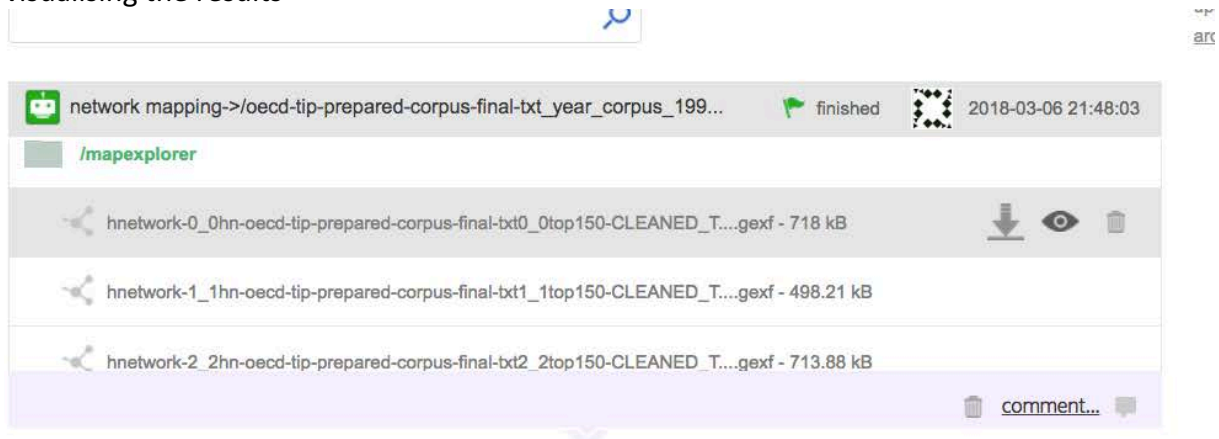
☐ growing

Click on “Start script”

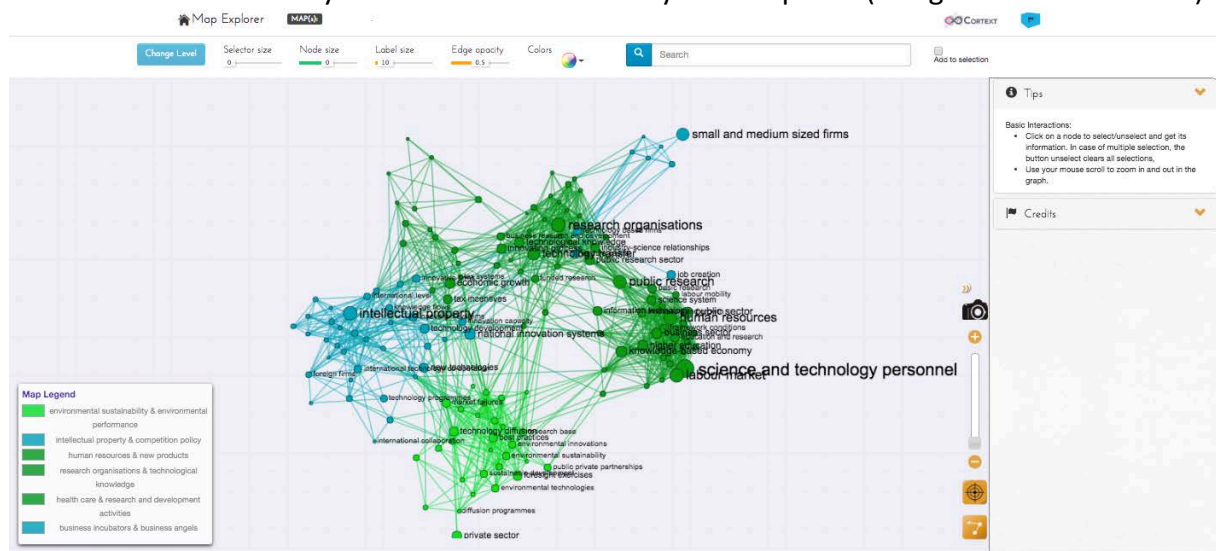
start script

## V) For visualising the maps

Click on “mapexplorer”: you have the 3 maps (one by time slice). Click on the “Eye” icon for visualising the results

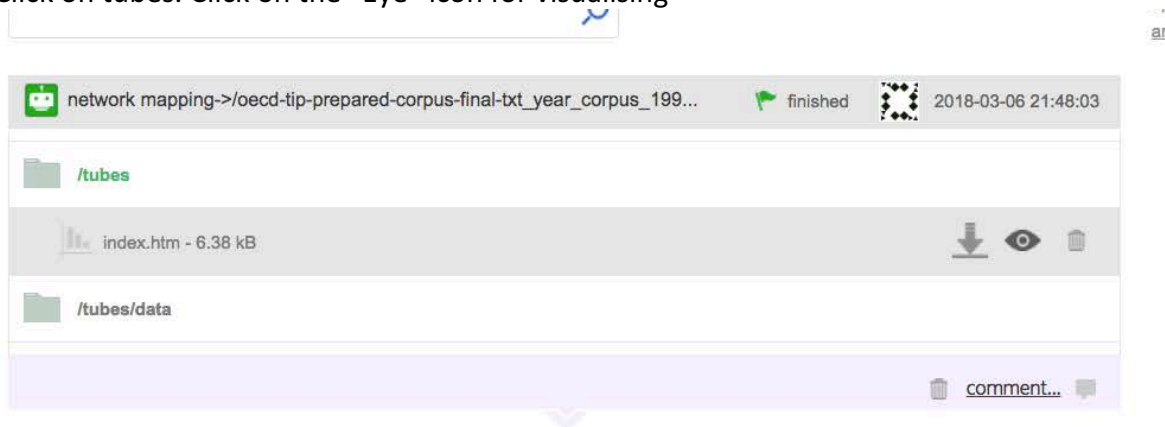


You visualise a network you can interact with and you can “print” (using the camera button)

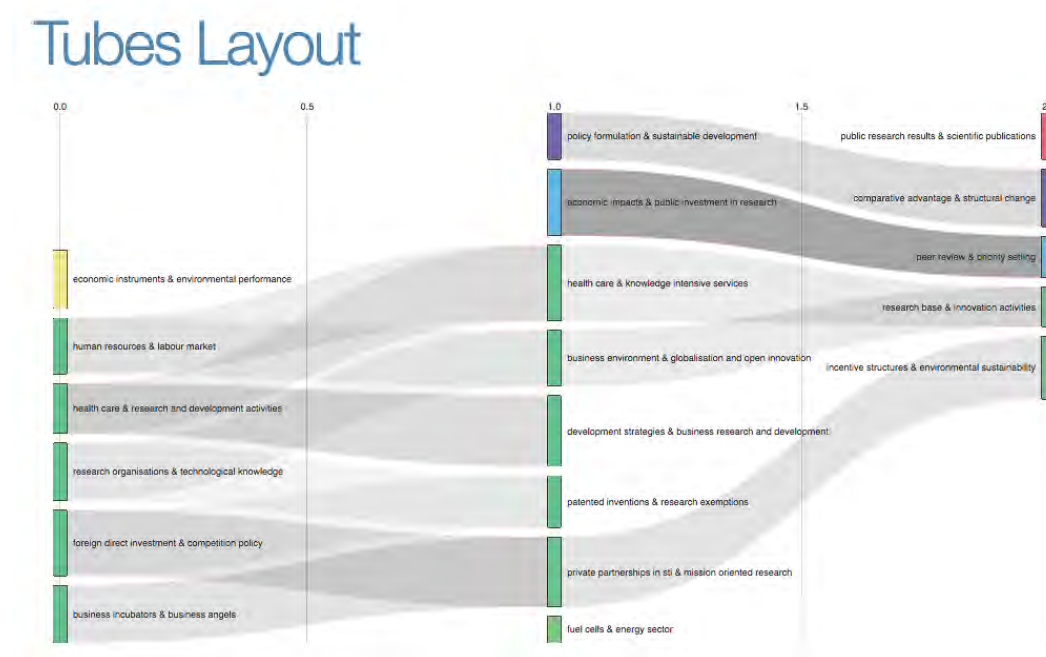


## VI) For connecting the maps across time periods

Click on tubes. Click on the “Eye” icon for visualising

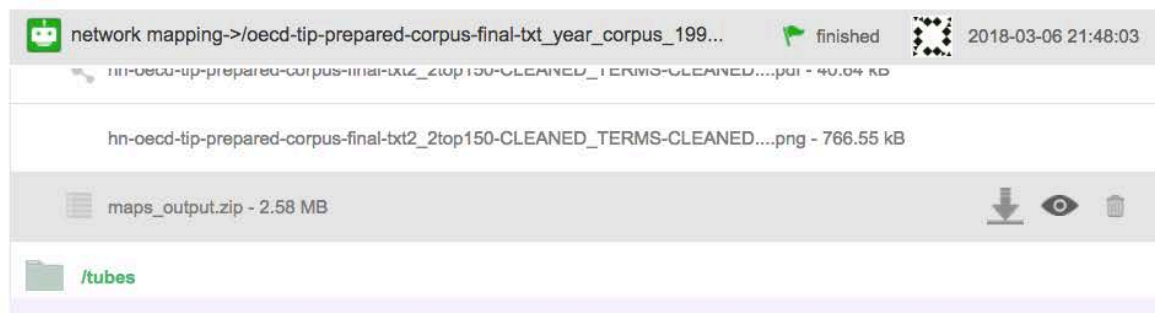


## Sankey Diagram linking the clusters across periods

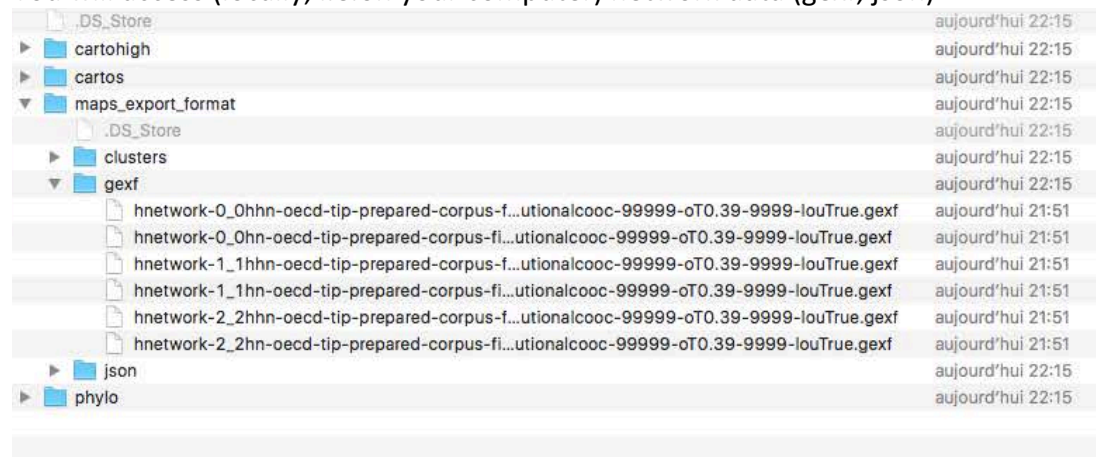


## VII) For exporting the network data

Click on “maps\_output.zip” and download, using the “descending arrow” icon.



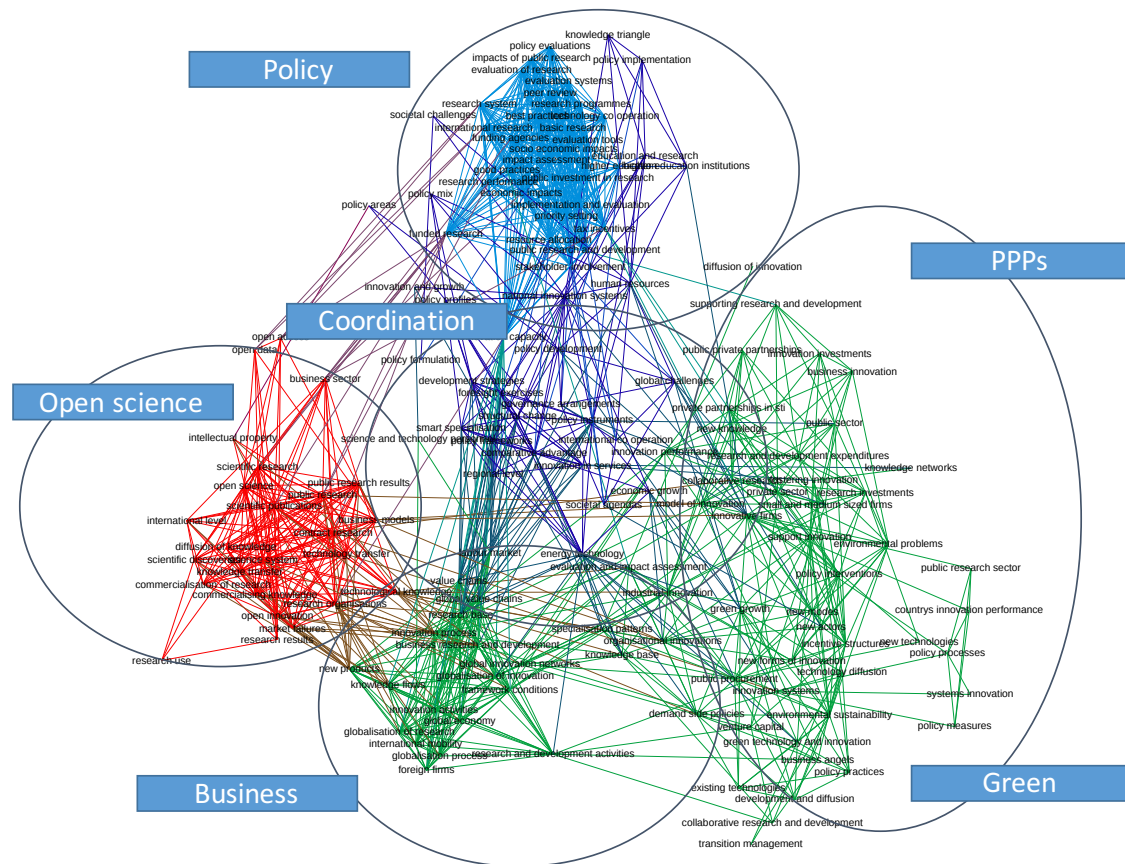
You will access (locally, i.e. on your computer) network data (gexf, json)





Then, you can work on this data, for instance, with the Gephi (open) software

**Third period (2009-2014) map, labelled after clusters' content expert analysis**



**Notes:**

- Help is available. You can thus access the CORTEXT MANAGER DOCUMENTATION [ps://docs.cortext.net/](https://docs.cortext.net/)
- You can collaborate on a project, inviting another Cortext registered user
- You can leave comments – interesting for collaborative uses