



*Gargantext :  
Collaborative Web Platform  
for Text-Mining  
<http://gargantext.org>*

May 18, 2015

# Sponsors



## Core Team



Mathieu Rodic

developer

[Website](#)



David Chavalarias

scientific board

[Mail](#)

[Website](#)



Alexandre Delanoë

project manager, scientific board,  
developer

[Mail](#)

[Website](#)



Elias Showk

developer

[Website](#)



Samuel Castillo J.

developer

[Mail](#)

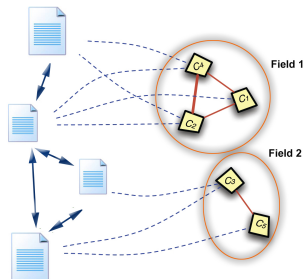
[Website](#)

# *Community*

- Mailing-lists:
  - User mailing-list: (soon)
  - Devel mailing-list: (soon)
- Interactive forum:
  - IRC: #gargantext channel on OFTC
- Code access:
  - Licence: AGPL
  - GIT access (soon)

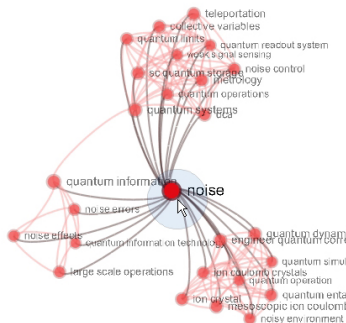
## *Interactive mapping of project portfolio*

- **Handle unstructured textual datasets** extract implicit relations (co-word analysis) through text-mining methods (no specific relation required, only some textual content).
- **Interactivity** Tools for real-time interaction with corpora. **Make your own idea** of what is there. Cumulative outcomes.
- **Easy-to-use & user-friendly** novices can play immediatly, become quickly confirmed users.



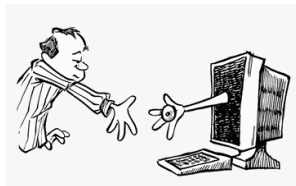
## *Interactive mapping of project portfolio*

- **Handle unstructured textual datasets** extract implicit relations (co-word analysis) through text-mining methods (no specific relation required, only some textual content).
- **Interactivity** Tools for real-time interaction with corpora. **Make your own idea** of what is there. Cumulative outcomes.
- **Easy-to-use & user-friendly** novices can play immediatly, become quickly confirmed users.



## *Interactive mapping of project portfolio*

- **Handle unstructured textual datasets** extract implicit relations (co-word analysis) through text-mining methods (no specific relation required, only some textual content).
- **Interactivity** Tools for real-time interaction with corpora. **Make your own idea** of what is there. Cumulative outcomes.
- **Easy-to-use & user-friendly** novices can play immediatly, become quickly confirmed users.



## *Mapping philosophy*

- Different views of the same complex network rather than THE map,
- Most valuable information about complex networks is obtained by map manipulations and queries,
- At any moment the user should be able to go from the aggregated level (the graphs) to the micro level (the documents/Ngrams selection) and back.



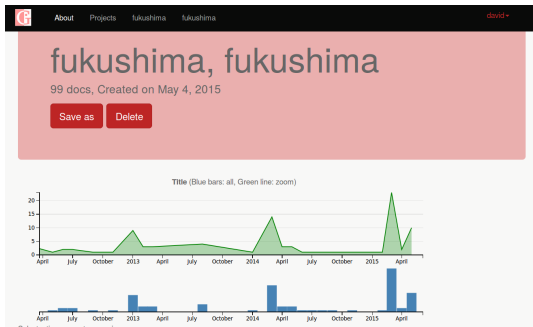
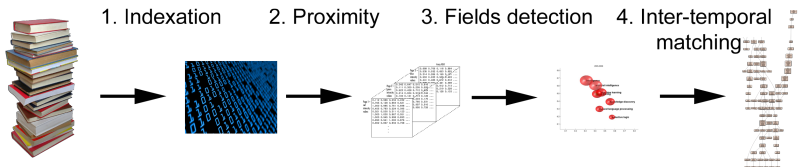
## *Mapping philosophy*

- Different views of the same complex network rather than THE map,
- Most valuable information about complex networks is obtained by map manipulations and queries,
- At any moment the user should be able to go from the aggregated level (the graphs) to the micro level (the documents/Ngrams selection) and back.

## *Mapping philosophy*

- Different views of the same complex network rather than THE map,
- Most valuable information about complex networks is obtained by map manipulations and queries,
- At any moment the user should be able to go from the aggregated level (the graphs) to the micro level (the documents/Ngrams selection) and back.

# Mapping projects with Gargantext



## *Gargantext is Database agnostic*

Custom parser for each database:

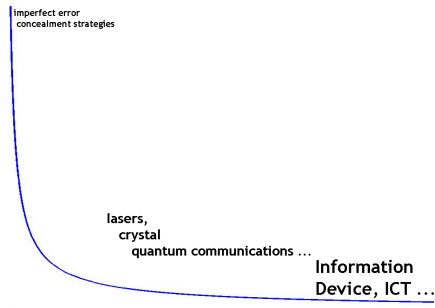
- Pubmed (XML format and direct queries)
- Web Of Science (ISI Format)
- Scopus (RIS Format)
- Europress (HTML Format)
- Zotero (RIS Format)
- Jstor (RIS Format)

## *Main views*

1. Global view of the corpus to clean the corpus
2. Advanced charts to explore the corpus
3. Explorer to interact with the graphs
4. Document view to read and annotate.

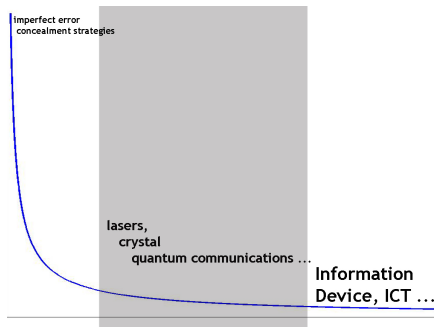
## *Some notion about terms selection*

- Powerlaw distribution is frequent
- Most interesting terms for maps are those with medium range occurrences.
- Each category of terms will bring characteristic patterns to the maps.



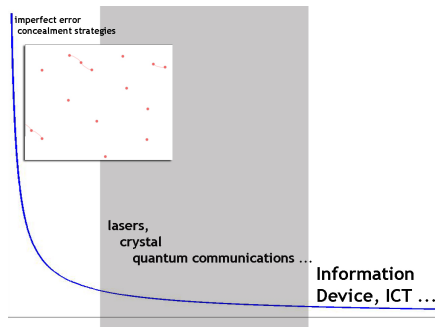
## *Some notion about terms selection*

- Powerlaw distribution is frequent
- Most interesting terms for maps are those with medium range occurrences.
- Each category of terms will bring characteristic patterns to the maps.



## *Some notion about terms selection*

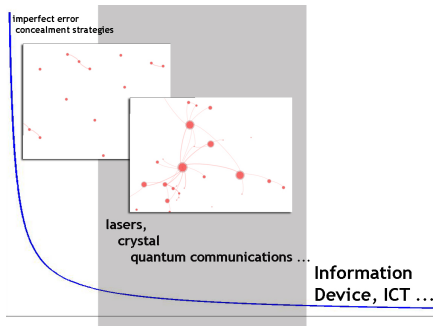
- Powerlaw distribution is frequent
- Most interesting terms for maps are those with medium range occurrences.
- Each category of terms will bring characteristic patterns to the maps.





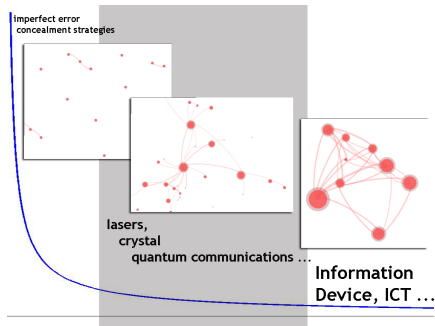
## *Some notion about terms selection*

- Powerlaw distribution is frequent
- Most interesting terms for maps are those with medium range occurrences.
- Each category of terms will bring characteristic patterns to the maps.



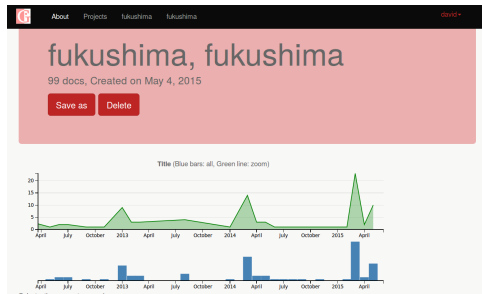
## *Some notion about terms selection*

- Powerlaw distribution is frequent
- Most interesting terms for maps are those with medium range occurrences.
- Each category of terms will bring characteristic patterns to the maps.



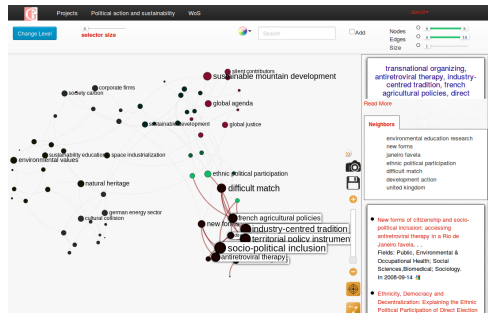
## About Gargantext

- **Gargantext** is a software for the production, the exploration and annotation of projects maps. It includes text-mining and natural language processing technologies, reconstruction methods of thematic landscape and visualisation tools.



# About Gargantext

- **Explorer** is a stand alone interface for the exploration of projects maps created by Gargantext. It can be ran on a personal computer or online.



## *Some basic operators*

- **Search** for a string
- Select one or several nodes,
- Switch to the other bipartite component,
- Change level from global view to neighborhood view,
- Adjust filter Unfold the graph removing edges and nodes



## *Some basic operators*

- **Search** for a string
- **Select** one or several nodes,
- **Switch** to the other bipartite component,
- **Change level** from global view to neighborhood view,
- **Adjust filter** Unfold the graph removing edges and nodes



## *Some basic operators*

- **Search** for a string
- **Select** one or several nodes,
- **Switch** to the other bipartite component,
- **Change level** from global view to neighborhood view,
- **Adjust filter** Unfold the graph removing edges and nodes



## *Some basic operators*

- **Search** for a string
- **Select** one or several nodes,
- **Switch** to the other bipartite component,
- **Change level** from global view to neighborhood view,
- **Adjust filter** Unfold the graph removing edges and nodes





## *Some basic operators*

- **Search** for a string
- **Select** one or several nodes,
- **Switch** to the other bipartite component,
- **Change level** from global view to neighborhood view,
- **Adjust filter** Unfold the graph removing edges and nodes



## Queries examples

- Find the core/periphery of the project network or the semantic network



- Map the semantic background of a part of the portfolio



- Target a specific domain



- Find the semantic and project background of an expression



- Find a path between two projects or terms



## Queries examples

- Find the core/periphery of the project network or the semantic network



- Map the semantic background of a part of the portfolio



- Target a specific domain



- Find the semantic and project background of an expression



- Find a path between two projects or terms



## Queries examples

- Find the core/periphery of the project network or the semantic network



- Map the semantic background of a part of the portfolio



- Target a specific domain



- Find the semantic and project background of an expression



- Find a path between two projects or terms



## Queries examples

- Find the core/periphery of the project network or the semantic network



- Map the semantic background of a part of the portfolio



- Target a specific domain



- Find the semantic and project background of an expression



- Find a path between two projects or terms



## Queries examples

- Find the core/periphery of the project network or the semantic network



- Map the semantic background of a part of the portfolio



- Target a specific domain



- Find the semantic and project background of an expression



- Find a path between two projects or terms



# *Perspectives*

- **Miam—Stop list—words manager**
- Multipartite heterogeneous network : still simple and reactive
- Dynamics: Phylogenies landscapes

## *Perspectives*

- **Miam—Stop list—words manager**
- **Multipartite heterogeneous network** : still simple and reactive
- **Dynamics**: Phylogenies landscapes



# *Perspectives*

- **Miam—Stop list—words manager**
- **Multipartite heterogeneous network** : still simple and reactive
- **Dynamics**: Phylogenies landscapes