

OPPORTUNITÉS DU TDM POUR L'ANALYSE DE L'ÉVOLUTION DES SCIENCES

David Chavalarias

Directeur de Recherche CNRS

EHESS/CAMS & Institut des Systèmes Complexes de Paris IdF

<http://chavalarias.com>

Journée INNOMÉTRIQUES 2016



QUESTION :

PEUT-ON PROPOSER UNE REPRÉSENTATION INTERACTIVE
DE L'ÉVOLUTION DES SCIENCES À PARTIR DE L'ANALYSE
DE SES PRODUCTIONS ?

Enjeux

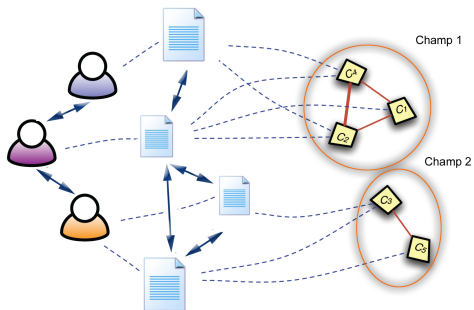
- > 1,5M d'articles publiés par an
- 8-9% de croissance/an de la production
- Principale source de connaissances spécialisées avec un focus amont

Conséquences :

- Il est de plus en plus difficile de faire un état de l'art : perte de temps et risque de redondance des recherches,
- Il est de plus en plus difficile de détecter les nouvelles opportunités de recherche et les innovations scientifiques,
- Il est de plus en plus difficile d'avoir une vision claire du développement d'un domaine, impact sur le pilotage de la recherche.

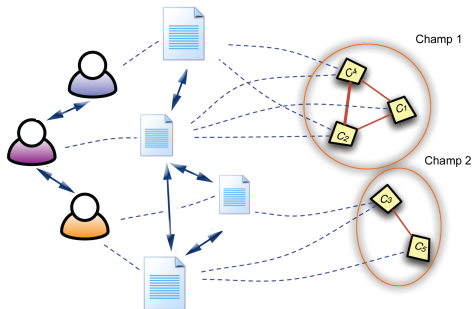
Structures multi-partites

Les données et méta-données des publications scientifiques permettent d'extraire les réseaux multi-partite de la science.



Structures multi-partites

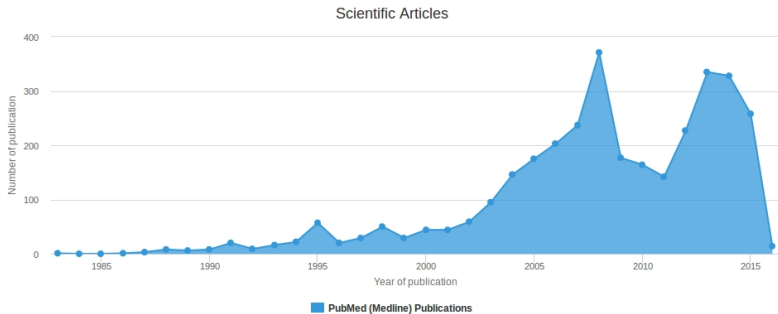
Les données et méta-données des publications scientifiques permettent d'extraire les réseaux multi-partite de la science.



Logiciels CNRS libres et open source de l'ISC-PIF :
<http://gargantext.iscpif.fr> & <http://multivac.iscpif.fr>

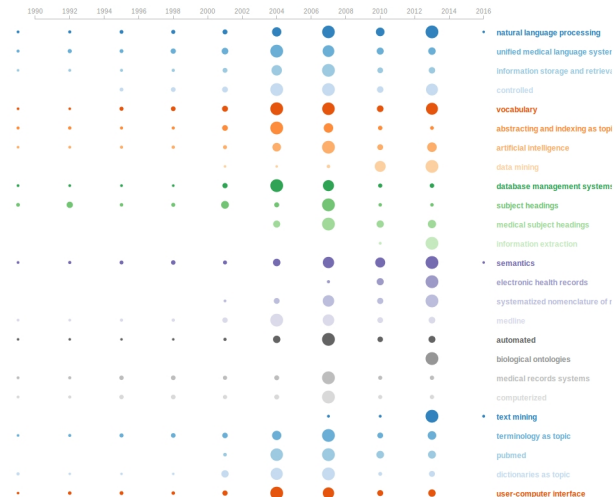


Evolution des sciences



Evolution des mentions de *Natural Language Processing* dans Pubmed
(source : Multivac)

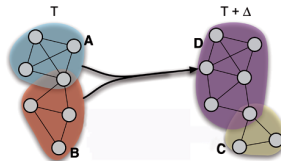
Micro-dynamiques : évolution des voisinages



Evolution du voisinage de *Natural Language Processing* (source : *Multivac*)

Structure multi-échelle des thématiques scientifiques

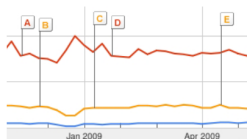
Macro Phylomemy



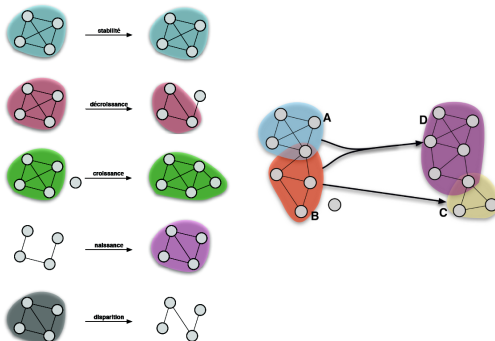
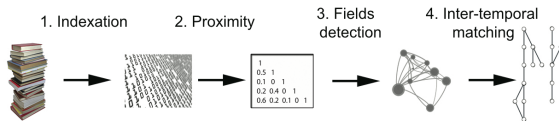
Meso clusters



Micro terms



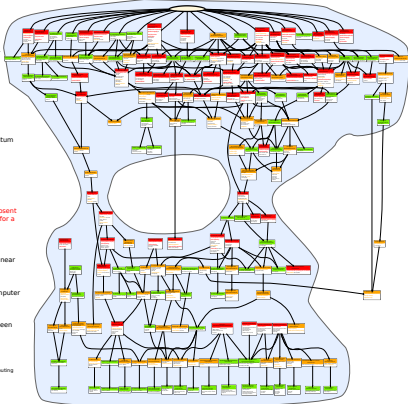
Dynamique des sujets (niveau "Meso")



Phylomemie des sciences (Chavalarias 2016 - HDR)

ORDINATEURS QUANTIQUES

- 1994 - Peter Shor algorithm to factorize large integers
- First schemes for quantum error correction.
- First realization of a quantum logic gate
- 1996 - First quantum database search algorithm.
- First public call for research proposals in quantum information processing (US Gov. & Army).
- 1998 - First experimental demonstration of a quantum algorithm.
- First working 3-qubit NMR computer
- 2001 **Negative result**
- Demonstration by Noah Linden and Sandu Popescu that entanglement (so far absent from experiments) is a necessary condition for a large class of quantum protocols.
- First execution of Shor's algorithm
- 2002 Quantum computation roadmap.
- 2003 - Quantum controlled-not gates using only linear optical elements
- DARPA Quantum Network operational
- 2004 - First working pure state NMR quantum computer
- 2005 - First quantum byte, or qubyte
- First transfer of quantum information between "quantum memories"
- 2006 - 2007 acceleration of discoveries.
Cf. https://en.wikipedia.org/wiki/Timeline_of_quantum_computing



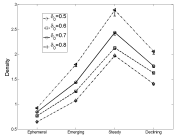
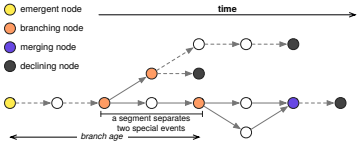
LEGENDE

■ Nombre de documents du WoS pour la requête ("quantum communication" OR "quantum information" OR "quantum computing" OR "quantum computer" OR "quantum processing" OR "quantum algorithm")

■ Première apparition du terme dans la phylomemie



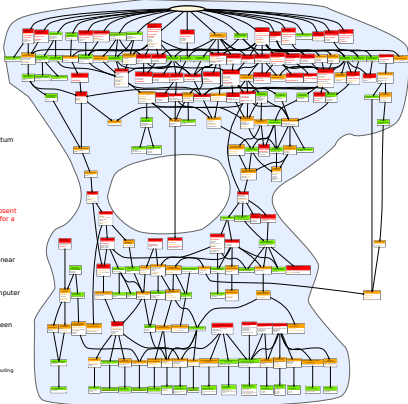
- 1999 **Membrane computing, natural**
- 2001 **EUV, Quantum matter**
- 2002 **Quantum memory channel, topological quantum computing, magnetodielectric**
- 2003 **One-way quantum computing, superconducting nanocircuits, pulse engineering, cluster states**



Phylomemie des sciences (Chavalarias 2016 - HDR)

ORDINATEURS QUANTIQUES

- 1994 - Peter Shor algorithm to factorize large integers
- 1995 - First schemes for quantum error correction.
- First realization of a quantum logic gate
- 1996 - First quantum database search algorithm.
- First public call for research proposals in quantum information processing (US Gov. & Army).
- 1998 - First experimental demonstration of a quantum algorithm.
- First working 3-qubit NMR computer
- 2001 **Negative result**
Demonstration by Noah Linden and Sandu Popescu that entanglement (so far absent from experiments) is a necessary condition for a large class of quantum protocols.
- First execution of Shor's algorithm
Quantum computation roadmap.
- 2002 - Quantum controlled-not gates using only linear optical elements
- DARPA Quantum Network operational
- 2004 - First working pure state NMR quantum computer
- 2005 - First quantum byte, or qubyte
- First transfer of quantum information between "quantum memories"
- 2006 - 2007 acceleration of discoveries.
Cf. https://en.wikipedia.org/wiki/Timeline_of_quantum_computing

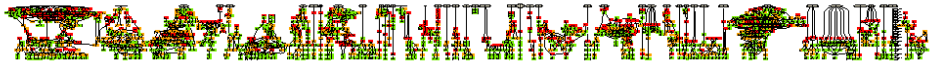


LEGENDE

■ Nombre de documents du WoS pour la requête ("quantum communication" OR "quantum information" OR "quantum computing" OR "quantum computer" OR "quantum processing" OR "quantum algorithm")

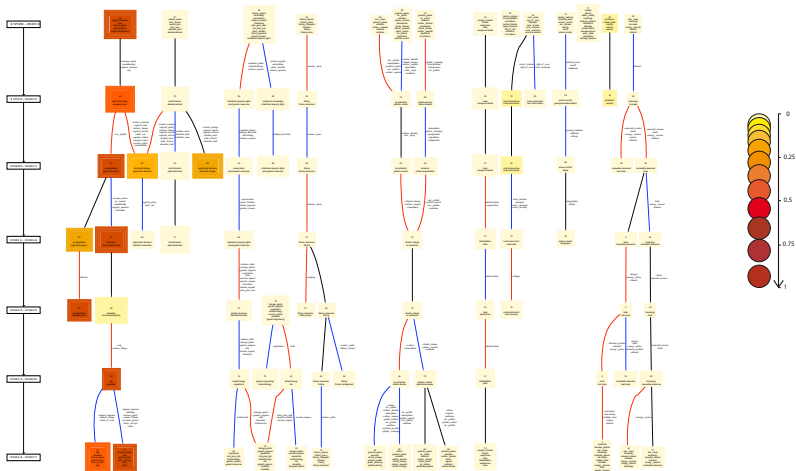
■ Première apparition du terme dans la phylométrie

- 1999 **Membrane computing, natural**
- 2001 **EUV, Quantum matter**
- 2002 **Quantum memory channel, topological quantum computing, magnetodielectric**
- 2003 **One-way quantum computing, superconducting nanocircuits, pulse engineering, cluster states**



Evolution des acteurs dans le paysage scientifique

- Contextualiser les parcours des chercheurs ... **qui contribuent à l'évolution du paysage**



Analyser l'évolution des sciences

- **Contextualiser** l'analyse des productions scientifiques
- Qualifier la diversité des profils de chercheurs
- La science en train de se faire. Qualifier les différentes branches de la science : émergence, "à la mode", mature, etc.

Analyser l'évolution des sciences

- **Contextualiser** l'analyse des productions scientifiques
- **Qualifier la diversité des profils** de chercheurs
- **La science en train de se faire.** Qualifier les différentes branches de la science : émergence, "à la mode", mature, etc.

Analyser l'évolution des sciences

- **Contextualiser** l'analyse des productions scientifiques
- **Qualifier la diversité des profils** de chercheurs
- **La science en train de se faire.** Qualifier les différentes branches de la science : émergence, “à la mode”, mature, etc.

Analyser l'évolution des sciences

- **Contextualiser** l'analyse des productions scientifiques
- **Qualifier la diversité des profils** de chercheurs
- **La science en train de se faire.** Qualifier les différentes branches de la science : émergence, “à la mode”, mature, etc.

Logiciels CNRS libres et open source de l'ISC-PIF :
<http://gargantext.org> & <http://iscpif.fr>



Conclusions

Il devient nécessaire de trouver de nouvelles manières d'interagir avec les productions scientifiques (et les productions textuelles en général).

Ces nouvelles pratiques vont avoir un impact considérable sur les choix stratégiques individuels des chercheurs et le pilotage de la recherche et doivent être régulées :

- → **Open software.** Ces pratiques doivent s'appuyer sur des outils transparents comme les logiciels libres,
- → **Open data.** Elles doivent s'appuyer sur les données les plus exhaustives possibles pour minimiser les biais dans l'analyse des productions : méta-données ouvertes pour l'ensemble de la production, possibilité de faire du TDM sur le texte plein et d'en communiquer les résultats agrégés.

L'accès de la communauté de recherche à une analyse statistique et aux méta-données de ses propres productions est une forme légitime d'empowerment.

MERCI POUR VOTRE ATTENTION



David CHAVALARIAS

Web : <http://chavalarias.com>

In collaboration with : Alexandre Delanoë, Mazyar Panahi, Romain Loth, Samuel Castillo, Jean-Philippe Cointet.



Rendez-nous visite
iscpif.fr