

Characterising the Localisation of Projects Collaborations in Research Dynamics: methodological requirements and results for new visualisations of heterogeneous networks

Tari T.^a, Caron P.^a, Breucker P.^{a,b}, Barbier M.^a

a. INRA SenS, Université Paris-Est

b. IFRIS CorTexT, Université Paris-Est

Research subject and hypothesis

This communication proposes to discuss the construction of methodological requirements on databases building and software development, and aspires to show some concrete results in visualising heterogeneous networks of research dynamics considered through projects ecology.

Our reflection is grounded in the growing needs, either for decision makers or researchers of the STS and SPS communities to relay their analysis of facts on a convenient visualisation of structural relationships between heterogeneous actants. Their configuration in dedicated databases is worthy to focus on as they reflect the endogenous dynamics of research and R&D activities. Our hypothesis is that the aims, perimeter, contents and selected projects of funding programmes represent a relevant account of the un-going technological and scientific dynamic on the one hand, and a relevant account of the mobilization and choices of scientific communities and science policy “makers” on the other hand. Those configurations rely firmly on spatial-based organizations, mixing European, national and regional scales in formal and informal clusters. Our perspective in the CorTexT Platform of IFRIS is to enrich the studies of sciences dynamics on customized databases of research and R&D projects that represent through territories performative associations of laboratories, scientific teams, R&D firms and lead-users.

Without ignoring the existence of a large array of scientific perspectives in Information Sciences about the measurement of science productions and science dynamics, we situate our work in the branch of analysis and visualisation of social networks. This field as well as indicators are central for evaluation and policy of science (Callon et al., 1986; Law et al., 1988). At present, the evolution of the analysis of scientific networks is largely attached to the question of characterizing collaborative and cognitive dynamics of knowledge production (Powell et al., 2005) and to the emergence of multi or trans-disciplinary emerging fields of research (Lucio-Arias, Leydesdorff, 2007) or paradigmatic field of research (Chavalarias, Cointet, 2008). Tracing and mapping knowledge in scientific database or in other electronic sources still represents a huge field of problems for many disciplines dealing with information. More locally, in relation to specific area of research, mapping heterogeneous networks appears to help the understanding of social dynamic of research activities (Cambrosio, Keating, Mogoutov, 2004; Cambrosio et al., 2006; Bourret et al., 2006).

Using co-word analysis tools (RéseauLu), we have already proposed a social study focused on regime of knowledge production in agricultural science and on the significance of sustainability (Barbier, Mogoutov et al., 2008). We identified two emergent yet lively research themes: biofuels and vegetal fibres, and realized specific bibliometrical studies on those subjects. We then devoted sociological studies based on heterogeneous sources to fibres (Caron et Barbier 2009) and biofuel & bioenergy research (Tari 2009). Bearing in mind this type of overallview on scientific knowledge we wanted to develop an approach on research projects in those domains.

Methodology

We decided to extend our scientometrical exploration to the ecology of projects. Unlike traditional scientific valorisations: articles, patents or even books, no global database gathers

information concerning this primordial way of researching and assembling the social in today sciences.

We wanted to question research dynamics through a large scope of sources related to research programs. Within those programs, the R&D project is our unit. The variety of project ecologies and organisational features lead us to conduct systematic data mining, data extraction and to constitute an adequate and robust information structure through the creation of a heterogeneous and relational database. Project properties, specifications concerning researchers, laboratories and institutions, as well as traditional scientometrical indicators but also geolocation data have been informed as precisely as possible. The relations between chosen entities are specified in order to formulate requests based on our research hypothesis. Either the results of hypertext annotation (on interviews), collaborative social networks or geolocation were then to be transformed in a matrix of non symmetrical relations. In order to visualise this kind of citation network in discourses about innovation and R&D relation, we mobilized the Visone software (Brandes, Raab and Wagner, 2001) and Geographic Information Systems (GIS) solutions. We developed a semi-automatic language translation tool between our databases, requests and various existing visualisation software.

The use of specific tools to index database and visualise networks in the context of characterising emerging domain of collaborations within Research and R&D projects represents a significant step forward to develop the field of scientometrics towards the characterisation of S&T Dynamic. It also enables to open the field of design of methodology and visualising solution to enhance new ways of tacking with relational data. This attitude toward methodological equipments for the visualisation of co-words clusters and mapping shares many ideas of shifting the use of tool from a scientific context to a science policy context (Noyons, 2001). We think that this turn should also go until the technical construction of facts within the scientific stance.

Tools in use (all open source software):

- Adequate and heterogeneous database construction: MySQL
- Visualisation of social networks : VISIONE, GEPHI
- Cartography based on geolocation: AMMAP, GRASS (GIS), SCAPETOAD.
- Languages for development: FLASH, GRAPHML, PHP, SQL, SHP, XML...

Existing and expected results

We propose to deliver a methodology of mapping research dynamics through R&D projects and a set of maps concerning actors of two domains: industrial fibres innovation and biofuel research. Our efforts enable the visualization of those networks according to our methodological tools and requirements. This geographical-based approach of actors and links is possibly to be enriched by a lexical analysis of a sociological qualitative survey, via a characterization of the main components of the strategic positions of entrepreneurs. Those materials allow interpretation on science dynamics, e.g. an analysis of the difficulties to innovate in the world of industrial hemp, the effectiveness of a scientific community constitution, an analysis of the new paradigm of public / private collaboration, the embeddedness of innovation clusters in regional landscape in bioenergies.

Assembling these methods into a unique grounded methodology represent a kind of “bricolage”, which also translates the hybrid nature of sociotechnical networks and the social dimension of entrepreneurship for innovation (Nohria and Eccles, 1991). We expect to associate very soon the mapping of geolocation to the mapping of research topic based on word-reference (Van Den Besselaar et al., 2006).

*Visualisations and mapping are available at:
http://www.cortext.fr/GeoLoc_Proj_BioFuelFib*

References

- Bar-Ilan J., 2008. Informetrics at the beginning of the 21st century—A review, *Journal of Informetrics*, 2:1–52
- Barbier M., François-Moll F., Vichnevskaja T. et Mogoutov A. 2008. Figures de la "soutenabilité" dans des bases de données scientifiques du "Web of Science". Exploration et discussion de traitements scientométriques, Communication au Séminaire Changement climatique, expertise et fabrications du futur : Modèles et scénarios, politique et gouvernance, Centre Koyré, Paris, 8 Avril 2008.
- Bourret P., Mogoutov A., Julian-Reynier C., and Cambrosio A., (2006). A New Clinical Collective for French Cancer Genetics A Heterogeneous Mapping Analysis, *Science, Technology, & Human Values*, 31 (4): 431-464
- Brandes U., Raab J., and Wagner D., 2001. Exploratory Network Visualization: Simultaneous Display of Actor Status and Connections. *Journal of Social Structure*, 2(4), 2001.
- Callon, M., J. Law, A. Rip (1986), *Mapping the Dynamics of Science and Technology*. London: The MacMillan Press Ltd.
- Cambrosio A., Keating P., Mercier S., Lewison G., and Mogoutov A., (2006). Mapping the emergence and development of translational cancer research, *European journal of cancer*, 24: 3140-3148
- Cambrosio A., Keating P., Mogoutov A. (2004). Mapping collaborative work and innovation in biomedicine: a computer assisted analysis of antibody reagent workshops, *Social Studies of Science*, 34 (3): 325-364.
- Caron P., Barbier M. (2009) Sustainable Innovation with Industrial Hemp and Flax in France: towards Non-Absorptive Intermediary Actors. Communication to the ERSA Conference, Territorial cohesion of Europe and integrative planning, Lodz.
- Chavalarias D, Cointet JP. (2008). Bottom-up scientific field detection for dynamical and hierarchical science mapping, methodology and case study, *Scientometrics*. 75(1): 37-50.
- Leydesdorff L, Hellsten I., (2006). Measuring the meaning of words in contexts: An automated analysis of controversies about 'Monarch butterflies,' 'Frankenfoods,' and 'stem cells', *Scientometrics*, 67 (2): 231-258.
- Lucio-Arias D, Leydesdorff L, (2007). Knowledge emergence in scientific communication: from "fullerenes" to "nanotubes", *Scientometrics*, 70 (3): 603-632
- Noyons E., (2001). Bibliometric mapping of science in a policy context, *Scientometrics*, 50(1): 83-98.
- Powell et al., 2005 W.W. Powell, D.R. White, K.W. Koput and J. Owen-Smith, Network dynamics and field evolution: the growth of interorganizational collaboration in the life sciences, *American*
- Tari T. (2009) A la recherche du biocarburant, entre promesses et controverses : étude de la forme projet dans les dynamiques de recherche sur les bioénergies en France. Mémoire de master EHESS Centre Alexandre Koyré, Paris.
- Van Den Besselaar P, Heimeriks G. 2006. Mapping research topics using word-reference co-occurrences: A method and an exploratory case study, *Scientometrics*, 68(3): 377-393.