How should we define, construct and compare clusters in emerging S&T fields? The case of nanotechnologies

Bernard Kahane[a], Aurélie Delemarle[a], Lionel Villard[a], Philippe Laredo[b]
[a]Université de Paris-Est, ESIEE Management, France
[b]Université de Paris-Est (ENPC) and University of Manchester (MBS)
b.kahane@esiee.fr

Introduction

Comparing potential and performance on a given S&T field is an exercice that public policy makers, companies executive and those in charge of research institutions have to perform. They need to assess their position and to relate them to those of others engaged in similar S&T fields. This is the equivalent of external or industrial analysis focused at the R&D stage. We will adress this issue through how work on nano science and technologies. Emerging fields constitute a specific challenge since that have no clear boundaries, are highly heterogeneous and that are evolving (for methodological developments on nano, see Mogoutov and Kahane 2007).

Classical levels of analysis are usually performed at country or institutional level which may both lead to strong perception and strategic bias. Indeed, countries who matter on R&D have very different sizes from continental level such as the US, China or Russia, to mid level such as France, Germany, UK, Korea down to small size such as Switzerland, The Netherlands, Israel, Singapore or Taiwan. This strongly impacts on relative potentials and outputs, thus impairing benchmarking for those who wish to understand which places are relevant and why. Further, distribution inside countries is not even with, if we take the case of nano S&T in France as an exemple, strong differentiation where areas such as Paris and Grenoble dwarf others. Meanwhile, analysis at institutional level would also provide a misleading view. Indeed, many areas are known to get their strength from the simulatenous presence of different types of institutions (universities, government labs, large companies, SMEs) which together provide an original and powerful potential. Others rely on the strong presence of many small and mid size companies. Meanwhile, in specfic cases, one single university or governement lab or company are the only relevant actor in a given area.

This has raised interest for analyses at a level which would be half way between countries and institutions. These approaches use either political-administrative delineation such as metropolitan areas in the US and NUTS regions in Europe, or the rather cloudly defined notion of clusters. The former approach is impaired by uneven definitions at world level which make difficult global analysis taking into account Asia, Europe and Americas as well as Oceania and even Africa. The latter one requires to use a bottom-up approach, choosing both the kind of data to build on and a methodology to achieve clustering.

Mid level clusters' analysis relevance is based in the litterature on the benefit from economies of localisation where agglomeration of activities take place (Krugman, 1991, and on S&T, Ricci 1999, Bonaccorsi 2005). Being located in a place allows most notably to benefit from "something in the air" (Marshall, 1890) as shown for instance by Saxenian (1994, 2007) in Route 128 and the silicon Valley, from spillovers (Grupp 1996, Feldman et al., 1999) or from the existence of an anchor tenant (Agrawal and Cockburn, 2003).

Advantages of localisation are however not constant and vary as the industry evolve: Mangematin et al. (2003), Powell et al. (2000) and Rothaermel (2008) have demonstrated that at the beginning of the biotechnology revolution, actors were looking for geographical proximity while as the industry matured, actors specialised along the value chain which allowed for a network like organisation not linked to a specific territory. Actors thus need to understand the dynamics of their industry to locate their R&D in an appropriate manner which require to combine both science and technology that are intrinsically interwoven. Here, we propose to distinguish between the dynamics of science and the dynamics of technology, thus relying on both publication and patent data for our analysis.

Methodology

Using a request previously described (Mogoutov and Kahane 2007), we extracted nanorelated data from Web of Science and Patstat databases. We then geographically locate both publications and patents. In our presentation, using nano S&T, we will first explain and discuss the bottom-up method we have developed to perform clusters' construction on a global scale. We will also describe ad hoc methodological design developed to deal with addresses ambiguities and holes which could affect localization at clusters' level. Dealing with patents confronts us to a specific challenge since contradictory results are obtained depending upon choices made between applications or (granted) patents, upon the selection of patent offices, the exclusion or inclusion of related applications or patents when one is extended to another area or to another claim (central data to define the kind of application or market that is at stake). We will report in our presentation on the choices we made and on the arguments on which those are based.

Results

In our work focused on nanotechnologies, we show, using only publications as a first step in our analysis, a concentration of potential and outputs on 200 clusters with strong uneven distribution at world level and strong hierarchies between clusters, and with very different actor composition of clusters (enabling us to discuss potential relations between the latter and growth patterns observed).

This hierarchy is challenged once patents data are introduced. Using the combination of publications and patents datta enable us to confront two dynamics of localisation and take a first measure of specialisation / agglomeration phenomena. Our provisional data highlight a strong focalisation effect where 50 clusters worldwide play a key role, offering a very specific geography that questions the direction of policies and raises issues about the relevant policy portfolio.

Thus, the presentation has two dimensions: one dealing with problems handling heterogeneous data sets and the possibilities offered by objective-driven solutions. The other deals with delineating the dynamics of new sciences and agglomeration phenomena showing as an illustration how this may impact policies.