Science-technology interactions: using NPLRs as *glue*

Jos Winnink winninkjj@cwts.leidenuniv.nl CWTS - Leiden University

Introduction

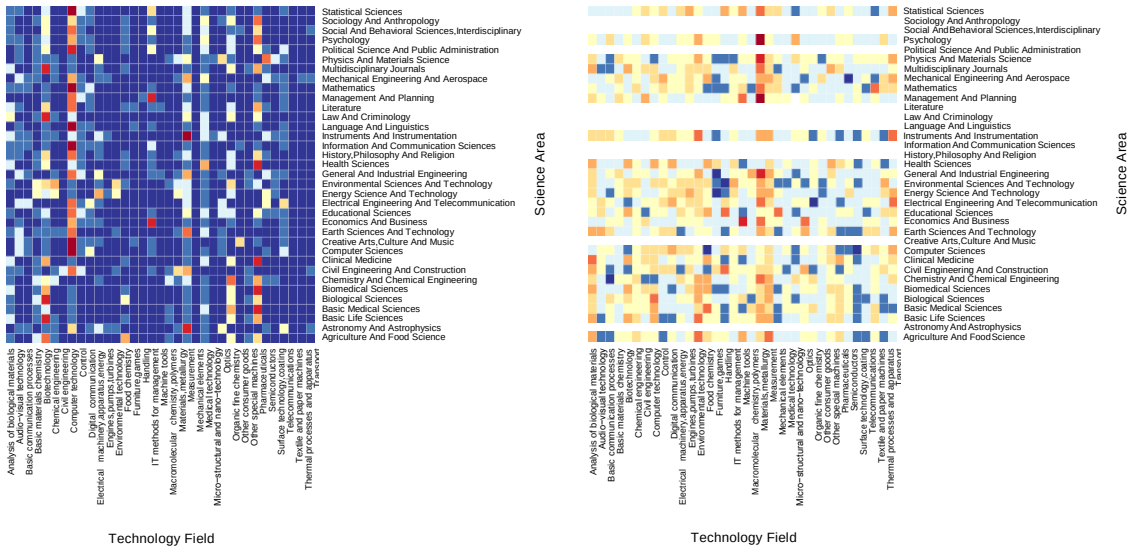
Scientific and scholarly research may result in a new discovery. The nature and impact of such a discovery on the cognitive structure and evolution of science may vary considerably. The impact of discoveries may extend beyond the domain of science and may be crucial steps towards technological applications, and to innovations and products. Scientific discoveries and their incorporation in technology are often interlinked in complex ways within research and development (R&D) systems. Such interactions may span several years, decades, or even centuries. The complex relations between scientific discoveries and technological developments has already — for dozens of years — been the subject of several studies. Some of the well-known landmark studies are the work conducted by Jewkes et al. (1958) and the *Hindsight* study (Isenson, 1969). The goal of these studies was not only to identify linkages between scientific discoveries and technological developments but also in finding relevant conditions that play an important role.

Methodology

Interactions between science and technology can be studied using various data sources. Patents are considered to represent technology and cite other patents as well as non-patent publications. Those citations in patents to non-patent publications are called non-patent-literature references (NPLRs). NPLRs that point to scholarly publications represent direct and visible linkages between science and technology. Due to the fact that comprehensive bibliographic databases for scientific publications and patents are nowadays available scholarly publications and patents can — in principle — be linked on the level of individual documents. Due to the fact that for NPLRs no prescribed format is used to store the citation information creating the linkages between science and technology on the basis of NPLRs used to be and still is a tedious (partial manual) task — even with the availability of digital libraries. In the *manual* era studies were therefore limited to case studies, with sometimes large numbers of cases. The availability of digital libraries opens up the possibility to analyse the influence of all science areas on all technology fields. By parsing and analysing the NPLR information stored in the PATSTAT database and matching the results with our in-house version of the WoS database we constructed an infrastructure that enables to establish linkages between science and technology on the basis of individual publications.

Some preliminary results and further research

By parsing the information describing the NPLRs and matching the results against a bibliographic database, in our case the Web of Science database, the interactions between science and technology can be established and studied on the level of individual documents. One possible option for the representation of the science-technology linkages is in the form of a *heatmap* in which science areas occupy one dimension and technology fields the other. The cells in such a science-technology matrix represent the *intensity* of the particular relation. As such a matrix contains information on the interaction of all science areas with all technology fields some of the combinations are unlikely to have a value that significantly differs from 0; e.g. we don't expect significant linkages for instance between



(a) Intensity of the science-technology linkages (b) Examiner citations vs. applicant citations

Figure 1: Interaction between science and technology based on NPLRs (1980-2017)

Arts and Humanities and one of the technology fields. A high-level picture based on 35 science areas, based on combinations of the 250+ WoS subject categories and 35 technology fields[1] a science-technology matrix can be constructed and is presented in two figures in Figure 1. The information is based NPLRs for all publications from 1980–2017 that could be matched. This science-technology matrix is shown in Figure 1a; the colour of a cell indicates the intensity of the relation. The colour scale runs from dark blue (low-intensity) via white and yellow to red (high-intensity).

Most active science-technology interaction can be found in the area of health science technologies (Biotechnology, Organic Fine Chemistry and Pharmaceuticals) and Computer Technology. Figure 1a also shows a red cell that represents the link between the science field 'Law and Criminology' with Biotechnology. Publication from the science area 'Astronomy and Astrophysics' have a relative large influence on 'Measurement Technology'.

In Figure 1b the difference in *perspective* on the science-technology interactions between the patent examiners and the patent applicants is visualised. This visualisation shows the data after double normalising the citation data for patent examiners and for patent applicants. A spot with a reddish colour indicates that the relation is stronger according to the examiner citations and a blueish colour that the link is seen weaker. White and Yellow spots indicate that examiners and applicants value this science-technology linkage more or less equal. Patent examiners and patent applicants have different roles in the patenting procedure; these different roles can and do lead to differences in the weights of the linkages of the science-technology relations as is illustrated in Figure 1b.

Further research will approach the science-technology linkages in at least the following directions:

(1) Figure 1b shows that the picture based on patent examiner cited NPLRs differs from the picture based on NPLRs cited by patent applicants. This raises the question if the picture based on all NPLRs should be preferred or that examiner citations or applicant citations should be preferred?

(2) the use of other and possibly more fine-grained classification systems of science and technology to be able to zoom in on particular relations;

[1]Information can be found in: "Concept of a Technology Classification for Country Comparisons" by Ulrich Schmoch, July 2008; http://www.wipo.int/edocs/mdocs/classifications/en/ipc_ce_41/ipc_ce_41_5-annex1.pdf

(3) adding the dimension of time to the analysis, so the evolution of the science-technology interaction can be analysed.

One of the research questions we like to answer is:

> *Is it possible to detect upcoming emerging technologies at early stage on the basis of the evolution of science-technology linkages?*

More and more detailed results will be presented at the conference.

References

Isenson, R. S. (1969). Project hindsight (final report). Technical Report AD495905, Office of the Director of Defense Research Engineering, Washington, DC, 20301.

Jewkes, J., Sawers, D., and Stillerman, R. (1958). *The Sources of Invention*. MacMillan, London.