Scientometric Mapping as a Strategic Intelligence Tool for the Governance of Emerging Technologies

Daniele Rotolo∗, Ismael Rafols†2,1, Michael Hopkins‡1, and Loet Leydesdorff§3

1SPRU — Science Policy Research Unit, University of Sussex, Brighton, United Kingdom
2Ingenio (CSIC-UPV), Universitat Politecnica de Valencia, Valencia, Spain
3Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, The Netherlands

Version: February 10, 2015

Abstract

This paper aims to provide evidence on how a set of recently developed scientometric techniques, namely overlay mapping, can function as a tool of ‘strategic intelligence’ to aid the governance of emerging technologies. We examine the potential these techniques have to inform, in a timely manner, analysts and decision-makers about relevant dynamics of technological emergence across three spaces: geographical, social, and cognitive. Our analysis relies on three case-studies of emerging technologies in the biomedical domain: RNA interference (RNAi), Human Papilloma Virus (HPV) testing technologies for cervical cancer, and Thiopurine Methyltransferase (TPMT) genetic testing. The case-studies are analysed and mapped longitudinally by using publication and patent data. Results illustrate the variety of ‘intelligence’ inputs overlay mapping can produce for the governance of emerging technologies and the flexibility and granularity — in terms of adaptability to different sources of data and selection of the levels of the analysis, respectively — these approaches confer to the investigation of emergence. These features also enable analysts and decision-makers to integrate and compare results from different contexts and cases, thus potentially functioning as platform for a ‘distributed’ strategic intelligence.

Keywords: scientometric overlay mapping; emerging technology; strategic intelligence; governance; case-study.

∗Corresponding author: d.rotolo@sussex.ac.uk, Phone: +44 1273 872980
†i.rafols@ingenio.upv.es
‡m.m.hopkins@sussex.ac.uk
§loet@leydesdorff.net
1 Introduction

An emerging technology can be defined as "a radically novel and relatively fast growing technology characterised by a certain degree of coherence persisting over time and with the potential to exert a considerable impact on the socio-economic domain(s) which is observed in terms of the composition of actors, institutions and patterns of interactions among those, along with the associated knowledge production processes" (Rotolo et al., 2015, p. 13). Emerging technologies have therefore the potential to change existing industries (Day and Schoemaker, 2000), the basis of competition (Hung and Chu, 2006), and human understanding or capabilities (Alexander et al., 2012). Their impact is also understood to be long term (typically with a 15-year horizon or so) (Porter et al., 2002), while their benefits are more likely to extend to the entire socio-economic system when they are of more 'generic' nature (Martin, 1995).

The status quo change nature of emerging technologies however "lies in the future and so in the emergence phase is still somewhat uncertain and ambiguous" (Rotolo et al., 2015, p. 13). Emerging technologies are not only associated with societal benefits, but also with unexpected, adverse, and destabilising changes (Collingridge, 1980). Their directionality is the result of a variety of factors including visions and expectations of the actors involved (e.g. Stirling, 2007; Van Lente and Rip, 1998). Emergence is of a reflective nature: actors are at the same time regulated by and regulating the emergence process.

For these reasons, the governance of emerging technologies has assumed an increasing relevance becoming priorities and part of the research agendas of many national governments. Developing policies to stimulate emergence towards societal benefit as well as to reduce the risk of the unintended effects or undesirable uses of emerging technologies is however a difficult activity. The governance of emerging technologies also assumes an emergent character resulting from the multitude of interactions of involved actors whereby the explicit attempts to shape arrangements are only one part of this process (e.g. Kooiman, 1993; Verbong and Geels, 2007). Governance is the result of both intentional and un-intentional influences, de facto including factors and actors non directly or purposefully aiming at governing lock-in and irreversibilities (David, 1985; Rip and Van Amerom, 2010). The risk of technological alternatives to be crowded out adds to this complexity (Martin, 1995).

In this context, the governing of emerging technologies requires tools of 'strategic intelligence', i.e. tools that are able to provide 'intelligence' inputs to the decision-making process for
the development of policy instruments capable to cope with the rapid, uncertain, and ambiguous evolvements of technological emergence (Kuhlmann et al. 1999; Stirling 2007). Conventional tools of strategic intelligence are science and technology foresight, innovation policy evaluation, and technology assessment. Despite the large variety of scientometric studies developing approaches for the detection and the analysis of emerging technologies (e.g. Glänzel and Thijs 2011; Porter and Detampel 1995; Small et al. 2014), frameworks on the use for strategic intelligence of scientometrics in the context of emerging technologies are less explored.

The present paper aims to provide evidence on how a recently developed stream of user-driven and 'interactive' scientometric mapping techniques, namely overlay mapping (Rafols et al. 2010),\(^1\) qualifies as a tool of strategic intelligence for the governance of emerging technologies. The key principle of these mapping approaches is to project data representing a focal subject area over a basemap, which represents the wider activity in science and technology in terms of publications and patents. The overlay can be constructed to represent the knowledge production of individuals, organisations, communities, or emerging fields. These techniques can be applied longitudinally to inform about the dynamics of emerging technologies across the geographical, social, and cognitive spaces of emergence as well as on combinations of those. For example, an organisation’s publishing and patenting activities can be projected on maps of science or Google maps across different levels such as journals, patent classes, or research topics (e.g. Klavans and Boyack 2009; Leydesdorff et al. 2014, 2013; Waltman and van Eck 2012).

We conduct our empirical analysis on three illustrative case-studies of emerging technologies in the biomedical domain. The case-studies are used as 'vignettes' to illustrate the 'intelligence' that overlay mapping approaches can potentially generate. The mapping on the selected case-studies is carried out since the early phases of emergence, and over decades of observation by using publication and patent data.

Results show the different types of perspectives on emergence that can be generated by analysts and decision-makers with relatively low efforts. These include urban areas producing the large amount of knowledge as well as the knowledge that is more likely to be of higher future impact, geographical distributions of key organisational actors in collaborative networks, and main scientific disciplines and technological domains involved in the emergence including dynamics of these. In addition, overlay mapping techniques show flexibility and granularity

\(^1\) The 'interactive' feature of the scientometric mapping enables the analyst or decision-maker to explore with details the considered data. Different areas of the maps can be detailed by zooming in and out.
in the analytical process. These tools are flexible in the sense that they are not constrained to the institutional boundaries of the databases where relevant data are maintained (Griliches, 1994). Perspectives on the emergence can be built by using same or similar search strings across multiple sources of data. Diverse granularity is provided by the possibility to cross multiple levels of analysis of the emergence process in order to reveal macro- and micro-dynamics. Flexibility and diverse granularity also favour the integration and comparison of results from different contexts which in turn supports the development of more ’distributed’ strategic intelligence on the emergence in science and technology. Overlay mapping may also suggest in a timely manner to analysts and decision-makers new directions of investigation as well as feeding into political discourse.

The paper is structured as follows. In the next section, we introduce the basic principle of overlay mapping approaches and the diverse possibilities of applications those techniques provide for the analysis of emerging technologies. We then introduce the three case-studies and data sources in Section 3. Results are presented in Section 4 and discussed in Section 5. Section 6 concludes the study.

2 Scientometric overlay mapping

The scientometric community has made great efforts in the development of a number of techniques to trace science and technology dynamics since the seminal works by Price (1965), Small (1973), Garfield (1979), and Callon et al. (1983). Well established examples of these methodological approaches are direct citation and co-citation analyses (e.g. Garfield et al., 1964; Small, 1977), bibliographic coupling (e.g. Kessler, 1963), and co-words analysis (e.g. Callon et al., 1983). A multitude of indicators, which mostly rely on publication and patent data, are based on this type of scientometric tools.

The growing attention to emerging technologies and their potential to change the status quo has also led scientometric scholars to focus their efforts on methodological approaches for the detection and analysis of emergence in science and technology (e.g. Glänzel and Thijs, 2011; Porter and Detampel, 1995; Small et al., 2014). Those efforts, jointly with the increase of computational power, the improved performance of processing algorithms, the higher technical

---

2 For a review of the scientometric contribution to the detection and analysis of emerging technologies see Rotolo et al. (2015).
capabilities in exploiting databases in a more comprehensive manner, have spurred the development of a number of mapping techniques among which overlay mapping (e.g. Rafols et al. 2010).

The overlay mapping approach allows rapid investigations across several dimensions and units of analysis of the emergence process. The basic idea underlying this approach is depicted in Figure 1. Publication and patent data gathered to represent an entity’s activity in science and technology are used to construct an overlay map. The entity can be an individual researcher (e.g. scientist, inventor), an organization (e.g. firm, university, research institute), a community, or, especially for the focus of this paper, an emerging field.

The overlay is then projected over a variety of basemaps. These include geographical maps, inter-organisational collaboration maps (co-authorship data), maps of science representing the global structure of science (as based on publication data) at multiple levels of analysis such as disciplines, journals, or topics, and maps of technological areas as identified by patent classes. These basemaps enable the tracing of emergence across various spaces. Here, we focus on geographical, social, and cognitive spaces. These were specified as the main dimensions in early scientometric projects (e.g. Narin 1976, Small and Garfield 1985) and related initiatives led by researchers in science policy and innovation studies (e.g. Irvine and Martin 1985), and patent analysis (e.g. Jaffe and Trajtenberg 2002).

Overlay mapping can complement the existing set of more conventional tools of strategic intelligence for the policy-making of emerging technologies. These mapping techniques can generate evolutionary perspectives to potentially reveal de facto (intentional and unintentional) government arrangements among factors and actors involved in the emergence process. These perspectives may facilitate timely analysis on relevant dynamics of emergence such as main geographical areas involved in the knowledge production processes and collaborative interactions among those, main actors in inter-organisational networks and their geographical distribution, domains of science and technologies involved in the emergence and associated dynamics across disciplines, journals, technological classes, and, in the case of biomedical technologies, medical areas. This, in turn, may support the development of policy instruments for governance that are more capable to cope with the rapid dynamics of emergence and associated uncertainty and ambiguity.

Flexibility in the use of diverse sources of data and multiple levels of granularity in the analysis are additional features that make overlay mappings as a useful tool for a 'distributed'
strategic intelligence on emerging technologies. Results from different contexts can reflexively and discursively be integrated and compared with relatively little additional effort. Basemaps provide common ground for comparisons. The choice of overlays and basemaps however depends on the specific questions the analyst or decision-maker aims to address.

To examine the use of overlay mapping as tools of strategic intelligence for the governance of emerge in science and technology, we first summarise a number of overlay mapping techniques that have been developed and reviewed in the domain of the information sciences and scientometrics. Specifically, we build on (i) Leydesdorff and Persson (2010) for mapping co-authorship relations in publications as overlay to Google Maps, (ii) Rafols et al. (2010) for mapping publications in terms of Web of Science (WoS) Categories, (iii) Bornmann and Leydesdorff (2011) for the mapping of highly-cited publications, (iv) Leydesdorff and Bornmann (2012) for the equivalent mapping of patents, and (v) Leydesdorff et al. (2012) for mapping medical innovations in terms of Medical Subject Headings (MeSH). Yet, differently from these studies, this paper goes

---

3 Reiterating the full technical details of these methods is not in the scope of this paper. The reader is referred to the above mentioned studies and corresponding webpages. In addition, it is worth noting that these methods build on a broader set of contributions on science and technology (S&T) mapping for policy use (e.g. Börner).
beyond the use of a single overlay mapping approach. It instead aims to provide an integrative synthesis of overlay mapping techniques for their use as strategic intelligence tool in the context of emerging technologies. To do so, these techniques are applied to three case-studies of emerging technologies.

3 Methods

3.1 Background on the case-studies

We conduct our analysis on three case-studies of emerging technologies in the biomedical domain: (i) RNA interference (RNAi), (ii) cervical cancer testing technologies for Human Papilloma Virus (HPV) and (iii) genetic testing for the Thiopurine Methyltransferase (TPMT) enzyme.

These case-studies were selected since they exhibit diversity in terms of their positions in the innovation chain. From a governance perspective, each technology raises different challenges: RNAi is a new technology with radical potential across the domains of therapeutics, diagnostics and as a basic research tool, but with highly fluctuating industrial interest; HPV testing is battling the entrenched technology of Pap testing, with powerful interest groups on both sides; TPMT testing technologies are exploited in a series of small clinical niches, mainly in developed countries where there are concerns that best practice may not be spreading. This diversity of case-studies provides us with the opportunity to discuss the use of overlay mapping to generate 'intelligence' inputs for the governance of emerging technologies in different contexts. However, at the same time, the three case-studies enable us to achieve a certain degree of external validity of the findings of our analyses (Gibbert et al., 2008; Yin, 2003). The comparison is somewhat controlled since these case-studies are from the biomedical domain and share primarily North American origin and similar periods of emergence, in the 1980s and 1990s. One would expect very different dynamics in other contexts such as electronic engineering — and possibly more difficulties in tracing the technologies with publications and patents. In the following, we provide background sketches for the case-studies to contextualise the scientometric maps presented later in the paper.

Firstly, RNAi is a molecular process that can silence the expression of genes. Gene regulation plays a critical role in the progression of cancers, genetic diseases, and infection agents. By
silencing specific genes one can stop the progression of a given disease. This ‘small RNA’ silencing mechanism was discovered in 1998. Its discovery reshaped the landscape of research on gene expression creating important expectations especially for the therapeutic applications (Leydesdorff and Rafols, 2011; Sung and Hopkins, 2006).

Secondly, HPV testing technologies are positioned within a specific domain of application, i.e. the detection for HPV virus, which is recognised as causing cervical cancer. Cervical cancer has a significant disease burden — about 500,000 new cervical cancers occur and cause about 250,000 deaths worldwide each year. This has led to the development of a large screening program, especially in developed countries. In the US more than 100 million tests are performed annually. While cervical cancer screening has been conducted for more than 50 years using the (cytology-based) Pap test to detect cancerous cells or cells potentially evolving into cancerous states, the discovery of the strong association between HPV infections (especially HPV types 16 and 18) and cervical cancer in the 1980s opened the space for the development of a competing and more sensitive testing technology based on molecular biology (Hogarth et al., 2012).

Thirdly, similarly to HPV test, TPMT testing technologies are positioned close to the applied-research domain. Yet, their application for clinical utility is contested across medical fields (e.g. different clinical guidelines supporting and discouraging the use of the test). TPMT testing is one of an emerging class of ‘pharmacogenetic tests’ which predict adverse events associated with pharmaceutical use (Hopkins et al., 2006). TPMT is an enzyme in the human body responsible for metabolising (i.e. breaking down) thiopurine drugs. Cytotoxic thiopurine drugs such as Azathioprine are used to treat a range of conditions including leukaemia, and autoimmune diseases (such as Lupus or rheumatoid arthritis). However, when a patient has mutations in the gene encoding TPMT, she/he is unable to metabolise the drug and is at increased risk of toxicity from a build-up of thiopurines. Following the discovery of genetic variations affecting drug metabolism, several types of TPMT testing technologies began to emerge in the 1990s with applications across a number of clinical fields of use such as oncology, dermatology,

---

4 Andrew Z. Fire and Craig C. Mello were awarded the 2006 Nobel Prize in ‘Physiology or Medicine’ 2006 for this discovery (Fire et al., 1998).
5 Harald zur Hausen, who discovered the association between the HPV and cervical cancer, was later awarded the 2008 Nobel Prize in ‘Physiology or Medicine’.
6 Cervical cancer remains however one of the most common cancers among women in developing countries that account for about 85% and 88% of new cases and deaths, respectively (GLOBOCAN 2008 available at http://globocan.iarc.fr).
7 “Health, United States, 2012 - With Special Feature on Emerging Care”, U.S. Department of Health and Human Services. Available at www.cdc.gov/nchs/data/hus/hus12.pdf.
gastroenterology, and rheumatology.

3.2 Data sources

Data were collected from multiple databases. We relied on Thomson-Reuter’s Web of Science (WoS) and MEDLINE/PubMed of the US National Library of Medicine (NLM) for publication data, and on the United States Patent and Trademark Office (USPTO) for patent data. These databases were queried with a set of keywords we identified by triangulating interviews with experts and quantitative research works on the case-studies (see Table 1). The interviews with experts in the fields of the selected case-studies are also used as lens of interpretations of the results we obtain from the application of overlay mapping approaches.

The retrieval of data for emerging technologies is challenging. Research topics often overlap and the vocabulary used to describe them is in flux (Robinson et al., 2007). Most up-to-date methods for retrieval rely on combining lexical search and clustering citation methods (e.g. Zitt and Bassecoulard, 2006), yet these approaches would be time-consuming for the non-expert analyst. We therefore deemed more suitable to use keyword-based searches given the parsimony and timeliness of the approach. We limited the search of keywords and their combinations in scientific articles’ titles since this approach tends to generate less false positives as compared to the searching in articles’ abstracts. Abstracts often contain technical and methodological terms not representing the core of knowledge the given article claims (Leydesdorff, 1989). An extended search may generally retrieve many additional records, yet with the risk of including many ones not closely related to the given emerging technology one aims to trace — it increases ‘recall’ at the expenses of losing ‘precision’.

The identification of patents that are relevant for the building of informative perspectives on emerging technologies requires a different approach. The incentives to patent are indeed different from those underlying the publication of scientific articles. The primary purpose of the patent system is to reward patentees by providing them a temporary monopoly to commercially exploit the patented inventions. This requires patentees to disclose the technical knowledge of the inventions. For this reason, valuable information is included in patent claims that inform on the scope of the technical knowledge of the considered invention (Hunt et al., 2007) — claims ”define the invention and are what aspects are legally enforceable” (USPTO Glossary).8 We

8 The USPTO Glossary is available at [www.uspto.gov/main/glossary](http://www.uspto.gov/main/glossary)
Table 1: Data sources and search strings.

| Case-study | Data   | Database       | Search string                                                                 |
|------------|--------|----------------|-------------------------------------------------------------------------------|
| RNAi       | Publications | ISI WoS       | TI=siRNA or TI=RNAi or TI="RNA interference" or TI="interference RNA"        |
|            |         | MEDLINE/PubMed | siRNA>Title or RNAi>Title or "RNA interference"[Title] or "interference RNA" [Title] |
| HPV testing| Publications | ISI WoS       | (TI=HPV* or TI="Human Papilloma Virus*" or TI="Human Papillomavirus*" or TI="Human Papilloma*virus*") and (TI=Cervical or TI=Cervix) and (TI=diagnos* or TI=test* or TI=assay or TI=detect* or TI=screen* or TI=predict*) |
|            |         | MEDLINE/PubMed | (HPV*[Title] or "Human Papilloma Virus*[Title] or "Human Papillomavirus*[Title]) and (Cervical[Title] or Cervix[Title]) and (diagnos*[Title] or test*[Title] or assay[Title] or detect*[Title] or screen*[Title] or predict*[Title]) |
|            |         | USPTO          | ACLM/(siRNA or RNAi or "RNA interference" or "interference RNA")              |
|            | Patents | USPTO          | ACLM/((HPV or "Human Papilloma Virus$" or "Human Papillomavirus$") and (Cervical or Cervix) and (diagnos$ or test$ or assay or detect$ or screen$ or predict$)) |
| TPMT testing| Publications | ISI WoS       | TI=TPMT or TI= "Thiopurine Methyltransferase"                                 |
|            |         | MEDLINE/PubMed | TPMT[Title] or "Thiopurine Methyltransferase"[Title]                           |
|            | Patents | USPTO          | ACLM/(TPMT or "Thiopurine Methyltransferase")                                |

*Source: authors’ elaboration based on interviews with experts and previous research works on the case-studies.*
therefore focused the search of keywords on patent claims. Issues related to the definition and
delineation of the boundaries of emerging technologies (e.g. the identification of an effective set
of keywords, limited data coverage) are further discussed in the next section.

The number of publications and patents that relate to each of the three cases, from 1982 to
2011, are reported in Table 2. While the first publications related to TPMT and HPV stem
from the early 1980s, data for RNAi appears only since 1998 when this silencing mechanism was
discovered and first published. Publication data from WoS and MEDLINE/PubMed allow for
relatively simple but informative reports on the emergence of these three technologies in terms
of published scientific articles. Relatively fast growth can be observed for the three cases. Yet, it
is clear that the pace of this growth as well as the scale of this emergence is significantly different
from one case to another in two respects. First, the growth in the number of publications for
RNAi is steeper than the other two case-studies. Second, RNAi and HPV testing technology
show an increasing number of publications for the entire observation period. Conversely, the
testing technology for TPMT enzyme seems to have reached its mature phase in the last few
years of observations.

Patent data reveal similar distinctive features. The production of patents (both granted
patents and patent applications) related to RNAi, for example, is relatively greater than in the
case of HPV and TPMT testing technologies. This is not surprising given the former’s wider
scope of potential applications. The declining of the patenting activity related to RNAi in the
last two years of observation is mainly due to the USPTO publishing patent applications after
18 months from the filing date. Although most patenting in RNAi occurs in small firms, the
decision of some large pharmaceutical companies, including Merck, Roche and Pfizer, to shut
down their R&D units on RNAi may have also resulted in a decreased interest on RNAi and
contributed to this decline (Lundin, 2011).

The patent application activity around HPV testing technology grows from 2002 to 2004
and then stabilises in the subsequent years with a peak of applications in 2009. The low number
of granted patents and patent applications for TPMT testing technology does not allow clear
trends in the production of technical knowledge to be deduced, other than to note this is a field
that appears to have attracted less attention.

---

9 Data on USPTO patent applications are accessible since year 2001. The filing year of patents was considered.
Table 2: Publication and patent data for the three case-studies.

| Year | RNAi ISI | RNAi MEDLINE | RNAi USPTO (granted) | RNAi USPTO (applications) | HPV testing ISI | HPV testing MEDLINE | HPV testing USPTO (granted) | HPV testing USPTO (applications) | TPMT testing ISI | TPMT testing MEDLINE | TPMT testing USPTO (granted) | TPMT testing USPTO (applications) |
|------|----------|--------------|----------------------|--------------------------|------------------|---------------------|-----------------------|-------------------------------|-----------------|------------------|-----------------------|-------------------------|
| 1982 | 0        | 0            | 0                    | 0                        | 3                | 2                   | 0                      | 0                             |                 |                  |                       |                         |
| 1983 | 2        | 1            | 0                    | 0                        | 2                | 0                   | 0                      | 0                             |                 |                  |                       |                         |
| 1984 | 0        | 0            | 0                    | 0                        | 0                | 2                   | 0                      | 0                             |                 |                  |                       |                         |
| 1985 | 4        | 2            | 0                    | 0                        | 3                | 2                   | 0                      | 0                             |                 |                  |                       |                         |
| 1986 | 12       | 10           | 0                    | 0                        | 5                | 1                   | 0                      | 0                             |                 |                  |                       |                         |
| 1987 | 22       | 14           | 4                    | 0                        | 5                | 4                   | 0                      | 0                             |                 |                  |                       |                         |
| 1988 | 17       | 16           | 0                    | 0                        | 2                | 0                   | 0                      | 0                             |                 |                  |                       |                         |
| 1989 | 30       | 24           | 3                    | 0                        | 1                | 1                   | 0                      | 0                             |                 |                  |                       |                         |
| 1990 | 32       | 28           | 0                    | 0                        | 2                | 2                   | 0                      | 0                             |                 |                  |                       |                         |
| 1991 | 36       | 31           | 1                    | 0                        | 6                | 5                   | 0                      | 0                             |                 |                  |                       |                         |
| 1992 | 41       | 49           | 2                    | 0                        | 5                | 5                   | 0                      | 0                             |                 |                  |                       |                         |
| 1993 | 18       | 27           | 1                    | 0                        | 11               | 9                   | 0                      | 0                             |                 |                  |                       |                         |
| 1994 | 29       | 30           | 4                    | 0                        | 8                | 5                   | 1                      | 0                             |                 |                  |                       |                         |
| 1995 | 24       | 28           | 5                    | 0                        | 17               | 14                  | 1                      | 0                             |                 |                  |                       |                         |
| 1996 | 35       | 32           | 1                    | 0                        | 11               | 8                   | 0                      | 0                             |                 |                  |                       |                         |
| 1997 | 33       | 28           | 2                    | 0                        | 15               | 10                  | 0                      | 0                             |                 |                  |                       |                         |
| 1998 | 3        | 2            | 1                    | 0                        | 37               | 33                  | 1                      | 0                             |                 |                  |                       |                         |
| 1999 | 15       | 10           | 1                    | 1                        | 41               | 35                  | 4                      | 0                             |                 |                  |                       |                         |
| 2000 | 42       | 33           | 4                    | 1                        | 43               | 41                  | 4                      | 0                             |                 |                  |                       |                         |
| 2001 | 56       | 52           | 5                    | 22                       | 50               | 52                  | 1                      | 5                             |                 |                  |                       |                         |
| 2002 | 166      | 126          | 25                   | 113                      | 54               | 43                  | 2                      | 13                            |                 |                  |                       |                         |
| 2003 | 427      | 280          | 45                   | 400                      | 69               | 59                  | 9                      | 29                            |                 |                  |                       |                         |
| 2004 | 785      | 523          | 83                   | 685                      | 59               | 59                  | 8                      | 19                            |                 |                  |                       |                         |
| 2005 | 892      | 658          | 86                   | 583                      | 105              | 95                  | 6                      | 21                            |                 |                  |                       |                         |
| 2006 | 932      | 782          | 124                  | 684                      | 94               | 86                  | 3                      | 24                            |                 |                  |                       |                         |
| 2007 | 1002     | 737          | 131                  | 752                      | 121              | 91                  | 6                      | 16                            |                 |                  |                       |                         |
| 2008 | 1032     | 827          | 106                  | 772                      | 134              | 106                 | 5                      | 21                            |                 |                  |                       |                         |
| 2009 | 1016     | 837          | 118                  | 767                      | 154              | 120                 | 5                      | 33                            |                 |                  |                       |                         |
| 2010 | 1029     | 904          | 89                   | 761                      | 121              | 99                  | 0                      | 18                            |                 |                  |                       |                         |
| 2011 | 1131     | 969          | 35                   | 410                      | 143              | 118                 | 0                      | 11                            |                 |                  |                       |                         |
| Total| 8528     | 6763         | 853                  | 5951                     | 1560             | 1357                | 77                     | 210                           | 540             | 350              | 16                    | 41                       |

*Source: authors' elaboration based on WoS, MEDLINE/PubMed and USPTO data that were collected in January 2013.*
3.3 Definition and delineations issues

A key preliminary input for the use of mapping and overlay techniques is the identification of a corpus — in this case, publications and patents associated with the emerging technology one aims to map. This task is generally much more problematic than one initially expects. The lack of a shared terminology due to the different meanings and expected uses associated with the technology makes the identification of a suitable set of keywords particularly challenging. Institutional vocabularies (e.g. MeSH terms or patent classifications) may also not be particularly helpful. Relevant categories (e.g. new terms or classes) may have not been yet included. This becomes even more problematic for more recent cases of emerging technologies.

The ambiguity of the definition of ’emerging technology’ (or more generally of ’technology’) adds to the difficulty of the task. A technology includes both the knowledge underpinning method or processes to fulfil a purpose (e.g. understanding of gene silencing) as well as the network of practices and components (e.g. the molecules and techniques that result in gene silencing). The case of RNAi is apt to highlight these issues.

RNAi is the naturally occurring process in which gene expression is reduced as a result of the breakdown of messenger RNAs. This process can be triggered by naturally occurring (endogenous) molecules, called microRNAs (miRNAs), or by externally inserted (exogenous) molecules, called small interfering RNAs (siRNAs). Soon after the discovery of RNA interference by Fire et al. [1998], siRNAs were recognised as valuable for gene silencing, both as tools for research and for therapeutic purposes. This led to a boom in public and private R&D investments [Haussecker 2008, 2012]. In parallel, but with some significant delay, it was realised that miRNAs were not a marginal phenomenon, but played a major role in gene regulation, including abnormal down or up regulation in certain diseases caused or earmarked by anomalous gene expression, such as many cancers.

In principle, then, to delineate the research boundary of RNAi one should include both siRNAs and miRNAs. However, it turns out that ’RNA interference’ as a technology (not as phenomenon) has become mainly known as the human-induced, exogenous interference, which was developed for therapeutic purposes. We then have two potential definitions of RNAi, one covering the entire field and one excluding miRNA — this latter, commonly adopted for the discussion of therapeutic applications.

Figure 2 illustrates that the observed dynamics are different. Whereas miRNA research is still booming, possibly due to their potential use as biomarkers in diseases such as cancers,
publications focused on siRNA have reached a plateau of about a thousand publications per year. The latter is possibly related to the challenges encountered in delivering siRNA in therapeutic applications that resulted in a retreat of pharmaceutical investment (Haussecker, 2012). These differences in trends support the views that the uses of miRNAs and siRNAs follow distinct trajectories, which it is best to differentiate. In this study, we focus on RNAi as research tool and for therapeutic applications, i.e. not including miRNA, which is not booming any more in terms of publications, but it is now beginning to reach applications — hence still emergent.

There is not any optimal solution to address the the fuzziness of the boundaries of emerging technologies. The analyst has often to carry out a process that may take several iterations and require a significant interpretative input from the experts of the considered technology before producing a suitable set of keywords for the corpus identification.

Figure 2: Delineating the boundary of RNAi. Scientific articles related to miRNA were retrieved by using the following search string in WoS: 'TI=microRNA*' or 'TI=miRNA*'. Source: authors’ elaboration.
4 Results

The use of overlay mapping approaches as tools of strategic intelligence for the policy-making process is discussed below. We specifically build perspectives on the emergence process of the three selected case-studies across the geographical, social, and cognitive spaces. The mapping approach enables the analyst to create syntheses of relatively complex information on the dynamics of emerging technologies and to visualise these in maps. The generated maps can be easily interpreted and analysed with visual inspection (e.g. emergence of new cities involved in the production of publications and patents, intensification of the collaboration activity, involvement of new scientific disciplines or technological areas). For this reason, the results described below are mostly illustrative and complemented with the qualitative understanding of case-studies.

However, it is worth noting that the aforementioned overlay mapping routines also allow the analyst to perform quantitative analyses. The overlay data are saved into a relational database, which can be queried to produce statistical information (for example, to perform hypothesis testing) on the considered emerging technology. Overlay data include: number of new countries/cities involved with the knowledge creation process, proportional increase of the number of publications in newly involved disciplines, and indexes of diversity of the scientific and technological areas involved. Due to space limitations, we discuss and report in the paper a sample of the results obtained — the entire set of maps and overlays is available as supplementary materials at www.interdisciplinaryscience.net/overlaymapping.

4.1 Tracing emergence in the geographical space

One can combine the use of overlays with geographical (base)maps to visualise, by using publications and patent data, the emergence process across cities, regions, and nations. Building on previous work (e.g. Bornmann and Leydesdorff 2011, Bornmann and Waltman 2011, Hu et al. 2012, Leydesdorff and Bornmann 2012), one can, for example, identify sites for a given emerging technology where highly-cited scientific articles were published more frequently than expected. This is depicted in terms of the sizes and colours of nodes.\footnote{The z-test for two independent proportions is used. The null hypothesis is the randomness in the selection of papers for a city (see Bornmann and Leydesdorff 2011). A threshold of top 10\% most-frequently cited scientific articles is selected.}

For the three case-studies of this article, we used overlays projecting publication data corre-
sponding to 5-year time windows. Using constant time windows across the case-studies allows us to capture the differing pace at which the different selected technologies have emerged in terms of (WoS) publications. Figure 3 depicts the results applied to HPV and TPMT testing technologies during the period 2002-2006.

Figure 3: Areas of highly-cited publications for the 2002-2006 period. Nodes are coloured dark green (red) when the difference between the observed number of top-cited publications and the expected one is positive (negative) and statistically significant ($p < 0.05$), light green (orange) otherwise. Source: authors’ elaboration.

For the case of HPV testing technologies, these maps identify, over the entire observation period, European urban areas producing highly-cited scientific knowledge more frequently than expected in London, Paris, and Amsterdam. New areas have also started to appear both in the north (nearby Copenhagen, Helsinki, and Jena) and south (nearby Barcelona, Bologna, and Turin) of Europe since the mid-1990s. The US sites contributing to this emerging technology with highly-cited articles are mainly located on the coasts, specifically in the area of Washington D.C., Baltimore, New York, and Boston, for the East Coast, and near San Francisco and San

The time window does not affect the results in this case. Similar dynamics are observed by using narrower or broader time windows (e.g. 3-year, 7-year). It is worth noting that in order to exploit all the available data up to the year 2011, the first time window for RNAi is 1998-2001, which includes four years of observations instead of five years. RNAi can indeed be observed starting from year 1998 when this gene silencing mechanism was discovered [Fire et al. 1998].
Diego for the West Coast. Georgetown University and the private company Digene in the area of Washington D.C. have played a key role for the development and the adoption of a test for the HPV detection in cervical cancer screening (Hogarth et al., 2012). The maps also reveal the rise of new sites producing highly-cited scientific knowledge in South America (e.g. near Sao Paulo, Buenos Aires) during the last ten years of observation.

This geographical mapping for TPMT testing technology locates, at the beginning of the observation period, highly-cited articles in the urban areas of Rochester and Memphis (US) as well as Sheffield (UK). However, as for HPV testing technology, new sites have started to appear since 1997 close to the East and West coasts (e.g. areas of Washington D.C., Boston, San Diego, San Francisco), across the UK (e.g. near London, Glasgow, Edinburgh) and Europe (e.g. Berlin, Madrid, Seville).

In the case of RNAi, the mapping shows the emergence of urban areas producing highly-cited articles on the East Coast of the US near Baltimore and New York, and, especially, in the area of Boston (see supplementary materials). The RNAi mechanisms were discovered by groups of scientists working in Massachusetts (Haussecker, 2008). However, while the maps in the subsequently years report the emergence of many sites all round the world, Massachusetts has persisted as the main area producing highly-cited scientific articles — this is revealed by the size of the node and as one of the hubs for RNAi biotech firms.

Interesting dynamics can be observed by combing these findings with those obtained using a similar approach but geo-localising patent data. These maps are reported for the 2002-2011 period in Figure 4 — the first time window (1998-2001) is excluded because there are only 11 patents, which constitute a too small sample for the statistical analysis. While the area of Denver has emerged as one of the many urban areas of high citation for the production of scientific knowledge in the later period of observation, these maps show this area producing highly-cited patents more frequently than expected. These patents related to the reagents used in RNAi (small interfering RNAs) while patents produced in centres nearby New York and Philadelphia related to therapeutic applications. The analysis does not identify any urban area of high citations in Europe or Asia.

This analysis shows how overlay mappings can provide the policy-making process with rele-

---

12 For this approach, we decreased the threshold from top 10% (as used for publication data) to the top 25% cited patents, necessary because the number of publications is an order of magnitude higher than the number of patents (Leydesdorff and Bornmann, 2012).

13 For this reason, these maps cannot be built for the HPV and TPMT testing technologies.
Figure 4: Areas of excellence (patent-based) for RNAi for the 2002-2011 period. Nodes are coloured dark green (red) when the difference between the observed number of top-cited patents and the expected one is positive (negative) and statistically significant ($p < 0.05$), light green (orange) otherwise. Source: authors’ elaboration.

vant inputs such as the list of geographical areas where main advancements for a given emerging technology were achieved and areas that are ‘unexpectedly’ listed. It may also inform on areas that persistently contribute over time to the development of the emerging technology as well as on new emerging ones. De facto coordination arrangements are also revealed. For example, the maps for HPV and TPMT testing technologies identify a number of urban areas in which highly-cited publications concentrate as located in developing countries (see supplementary materials). A further analysis of the collaboration networks (see below) revealed co-authorship links between these areas and others leading the advancements of these emerging technologies in the developed countries. These collaborations may have provided developing countries with the access to critical capabilities and resources to produce novel and high-quality knowhow, but perhaps indicating some crucial contribution (e.g. a genetic resource or patient population).
needed by researchers in developed countries too that was helpful.

It is important to note some limitations of the above-discussed approaches. Firstly, the geographical information reported in publication and patent data may not reflect the locations where the research was conducted. This limitation has to be taken into account especially when patent data are used. Inventors list their home addresses, which may not necessarily reflect the geographical locations where inventions were developed. Secondly, while the overlays built at city-level provide high granularity, they can represent sites located in the same urban area as two different nodes. For example, in the case of HPV testing technology, Silver Spring (a suburb of the US capital) was considered as a node different from Washington D.C.. These limitations suggest direction for future development of these techniques.

4.2 Tracing emergence in the social space

The structure of the collaborative relationships among the actors surrounding emerging technologies and their dynamics play a critical role in the emergence process (e.g. Latour, 1993; Powell, 1990). These connections are channels through which actors gain access to and mobilise knowledge, resources, and power. Networks of agents affect and are affected by emerging technologies, i.e. actors create social structures over the emergence that both enable and constrain their actions (Giddens, 1984). By using co-authorship data (e.g. Crane, 1972; Glänzel and Schubert, 2004; Wagner, 2008), the dynamics across this relevant space of emergence can be roughly traced — however, it should be noticed that many collaborations are not picked up by scientometric tools (Laudel, 2002). Novel techniques also allow building perspectives crosscutting both the social and geographical spaces (Leydesdorff and Rafols, 2011).

For example, Figure 5 shows the 2002-2006 collaboration networks at city-level for HPV and TPMT testing technologies. In this map, nodes are cities and the linkages between nodes are traced by using co-authorship data. Investigating the dynamics across these maps may provide informative perspectives that are derived by combining the geographical and social spaces. Examples of key empirical questions the overlay mapping approach may help to address in the policy-making process are: Where does an emerging technology arise? Does the collaboration network cluster in specific areas? Does the given emerging technology spread across cities, re-

14 Similar maps can be also built with the top-cited publications approach we described in the previous section. This provides additional perspectives on the structural position cities producing highly-cited knowledge occupy in the web of collaborative relationships (co-authorships). We make these maps available for the three case-studies as supplementary materials.
regions, and countries and, if yes, through which (collaboration) channels — which types of grant were important, which size of institutions, or types of knowledge brokers?

The collaborative network of HPV testing technologies, for example, discloses three relevant dynamics of the emergence process. First, a strong collaborative activity between the US (especially the areas of Washington and New York) and Europe (initially Germany) can be observed. This technology indeed started to emerge after a scientist, Harald zur Hausen, at German Cancer Research Centre proved HPV infections to be strongly associated with the development of the cervical cancer [zu Hausen 1987]. This discovery subsequently found important application in the US where extensive screening programs on cervical cancer were already in place. Here, a small biotech company, namely Digene, developed and marketed a series of FDA-approved HPV tests, which are recognised as significant milestones in the management of cervical cancer.

Second, the last ten years of observation show an increasing involvement of developing countries (e.g. Brazil, India) in the research networks. Cervical cancer in these countries is a significant social burden. HPV is sexually transmitted and the costs associated with screening the large at-risk population (a high proportion of adult women) are not always affordable. Third, globalisation of research on HPV testing technologies can be observed across the entire period as revealed by the density of the network of relationships across cities. The latter dynamic is also observed in the case of RNAi that diffuses all round the world, even more rapidly than HPV and TPMT testing technologies (see supplementary materials).

In the case of TPMT testing, further informative dynamics are revealed. The co-authorship network evolves starting from a strong collaboration between Rochester (US) and Sheffield (UK). This collaboration intensifies over the entire observation period while major collaborative networks within the UK and US national boundaries formed later, only since the 1990s. Subsequently, from 1997, we observed also the rise of the European network initially involving Germany, France, The Netherlands, and UK, and then including other countries such as Italy and Spain.

Building on these cross-cuttings of the geographical and social spaces, one can focus more attention on the social dynamics by looking at the structure of the web of relationships composing the network at a lower level of analysis such as the organisation-level. The network can be explored with algorithms that identify cohesive groups of organisations as well as public and private players occupying key positions. Social network analysis provides a broad range of measures that may support a quantitative assessment of the changes occurring over time in the
Figure 5: Crosscut on the geographical and social spaces of the emergence process for the 2002-2006 period. The size of nodes is proportional to the $\log_2$ of the number of scientific articles (plus one) that the given city (above)/organisation (below) published. We reported labels for the top 5% central organisations in the co-authorship network according to the degree centrality. Source: authors’ elaboration.
analysed network (Wassermann and Faust, 1994). It is worth noting that the interpretation of some of these measures is difficult because they are size dependent and networks featuring in emerging technologies show relatively fast growth.

For example, the lower part of Figure 5 depicts the inter-organisational networks corresponding to the aforementioned collaborative networks overlaid on the geographical space. While for HPV testing technologies a giant component of the network can be identified, the inter-organisational network for TPMT testing is highly fragmented, as revealed by the different separated groups of organisations (components), until the last five years of observation (Table 3). For clarity of the representation, Figure 5 depicts the largest component for HPV testing technology for the 2002-2006 period. We instead represented for TPMT testing technology components formed by at least three organisations.

The dynamics of the collaborative network surrounding the research activity on HPV testing technology show Digene occupying a strong and influential position within this network by collaborating with major institutions in the field of cervical cancer screening (e.g. National Cancer Institute, a public sector sponsor, and Kaiser Permanente, a large healthcare service provider and early test adopter). This eventually allowed Digene to influence the adoption of the test, for example by working with clinicians involved in the definition of medical guidelines (Hogarth et al., 2012). In other words, while Digene’s product development activity was regulated by the FDA for example, Digene was affecting the developments and dynamics in cervical cancer screening, too.

As discussed, the co-authorship network for TPMT testing technology is characterized by the presence of several separate components over a significant part of the observation period. The network seems to initially develop around the strong collaboration between Mayo Clinic (Rochester, US) and University of Sheffield and subsequently around three actors, who played a key role for the creation of a large network component in the later years of observation. These actors are St. Jude Children’s Research Hospital, University of Manchester, and Dr Margarete Fischer Bosch Institute. Further analysis of the content of publications and patent data revealed

Organisations’ names included in publication data of WoS present a number of variations, i.e. the same organisation may be spelled in different manners. We use The Vantage Point software to clean the data. This software specifically analyses and suggests groups of names that may refer to the same organisation by using a fuzzy algorithm that exploits also the information included in other fields of the publication data. We checked those suggestions for our sample of publications and confirmed those matches for which the manual desktop search over the Internet provided further support. Freeware routines for using institutional addresses but without this cleaning process can be retrieved at [www.leydesdorff.net/maps](http://www.leydesdorff.net/maps).
Table 3: Structural properties of the co-authorship network at organisation-level.

|                  | Time window | 1982-1986 | 1987-1991 | 1992-1996 | 1997-2001 | 2002-2006 | 2007-2011 |
|------------------|-------------|-----------|-----------|-----------|-----------|-----------|-----------|
| **HPV testing technology** |             |           |           |           |           |           |           |
| Nodes            |             | 20        | 130       | 173       | 265       | 471       | 816       |
| Ties             |             | 6         | 88        | 223       | 476       | 980       | 2075      |
| Density          |             | 0.03      | 0.01      | 0.01      | 0.01      | 0.01      | 0.01      |
| Average Path Length |           | 1         | 1.3       | 3.4       | 2.8       | 4.4       | 4.4       |
| Isolated (%)     |             | 8 (40%)   | 35 (27%)  | 43 (25%)  | 47 (18%)  | 55 (12%)  | 83 (10%)  |
| Components (minimum 3 nodes) |   | 0         | 15        | 11        | 16        | 23        | 40        |
| First largest component (%) | | 2 (10%)   | 7 (5%)    | 62 (36%)  | 83 (31%)  | 239 (51%) | 504 (62%) |
| Second largest component (%) | | 2 (10%)   | 6 (5%)    | 8 (5%)    | 12 (5%)   | 14 (3%)   | 9 (1%)    |
| Degree, Mean (Std.Dev.) | | 0.6 (0.5) | 1.3 (1.3) | 2.6 (3.0) | 3.6 (4.8) | 4.2 (4.7) | 5.1 (5.7) |
| **TPTM testing technology** | |           |           |           |           |           |           |
| Nodes            |             | 6         | 8         | 36        | 111       | 200       | 232       |
| Ties             |             | 5         | 3         | 28        | 123       | 203       | 413       |
| Density          |             | 0.33      | 0.11      | 0.04      | 0.02      | 0.01      | 0.02      |
| Average Path Length |           | 1.5       | 1.1       | 1.6       | 2.4       | 1.7       | 3.9       |
| Isolated (%)     |             | 1 (17%)   | 2 (25%)   | 7 (19%)   | 19 (17%)  | 43 (22%)  | 33 (14%)  |
| Components (minimum 3 nodes) |   | 1         | 0         | 3         | 11        | 22        | 19        |
| First largest component (%) | | 5 (83%)   | 2 (25%)   | 9 (25%)   | 25 (23%)  | 15 (8%)   | 82 (35%)  |
| Second largest component (%) | | 1 (17%)   | 2 (25%)   | 5 (14%)   | 17 (15%)  | 14 (7%)   | 9 (4%)    |
| Degree, Mean (Std.Dev.) | | 1.7 (1.4) | 0.7 (0.5) | 1.6 (1.4) | 2.2 (2.2) | 2.0 (2.0) | 3.6 (3.2) |

*Source: authors’ elaboration.*
St. Jude Children’s Research Hospital as conducting fundamental basic research in this domain as well as developing strong IP by holding key patents on genetic mutations used in TPMT testing, University of Manchester focusing on the issues related to the cost-effectiveness of using TPMT testing technology, while the Dr Margarete Fischer Bosch Institute — a large research foundation working on the customisation and improvement of drug therapy — hosted staff that held together a wider collaborative network through personal links (e.g. visiting fellowships) at other major centres in the field.\footnote{The dynamic collaborative networks of the three case-studies are also included as supplementary materials.}

The above discussed cases and techniques combined the geographical and social spaces of emergences and moved across units of analysis, in this case from the city- to the organisation-level. This provides a clear example of the flexibility and diverse granularity of the overlay mapping, thus providing intelligence on the constellations of actors involved in the emergence process, structure of the relationships among these actors, key actors and collaborations shaping the emergence as well as main channels where knowledge and resources may flow.

4.3 Tracing emergence in the cognitive space

As new technologies emerge, epistemic developments occur in terms of discoveries, novel theories, or changes in technical developments such as experimental systems, materials, methods, and instrumentation \cite{Joerges2002, Rheinberger1997}. These dynamics can be traced across the cognitive space by creating overlays of publications on basemaps of science that can be defined at different levels of analysis \cite{Klavans2009, Waltman2012}.

The publishing activity related to three case-studies can be, for example, projected across the map of science defined by the 225 WoS categories \cite{Rafols2010}. In this map, each node is a WoS category that can be assumed as proxy of a scientific subdiscipline. The projection (overlay) makes a node’s size proportional to the number of publications related to the given technology that were published in the given discipline the node represents. The different colours of nodes represent different clusters of disciplines. Figure\ref{fig:map} depicts the projections of publications related to the three case-studies. We report the map representing the structure of science (left) — the strength of each linkage is proportional to the similarity of citations patterns of two WoS categories — and the heatmap version (right), which shows relative density of publications in a
given area of the basemap. This combination provides an intuitive visualisation of the diffusion process of emerging technologies.

As for the previous analyses, we used overlays projecting the publishing activity according to 5-year time windows. While these maps show the rapid diffusion of RNAi technology across many disciplines such as molecular biology, oncology, biomedical research, and chemistry, the overlays of HPV and TPMT testing technologies reveal different directions of diffusion. HPV testing technology diffuses from basic research in oncology, pathology, and virology disciplines towards issues related to the public health. We interpret this dynamic as a representation of the extensive and ongoing debate on the practices adopted for the screening of the population. The debate has been focused on the suitability of HPV testing technology as first an adjunct to, and latterly a substitute for, the widely adopted Pap test (Hogarth et al., 2012).

TPMP testing technology diffuses from the basic research in pharmacology towards clinical disciplines such as gastroenterology and dermatology. Publication activity seems to equally spread in gastroenterology and dermatology disciplines during the 1992-1996 period. Yet, in the subsequent years, the volume of publications shrinks from dermatology area while continuing to grow in gastroenterology. This raises questions about the degree to which TPMT’s use has been contested in the different communities such as gastroenterologists, rheumatologists, and dermatologists.

It is worth noting that the techniques we have presented so far also suggest avenues of investigation an analyst may want to pursue for informing the policy-making process. For example, we followed the aforementioned dynamic by conducting an additional analysis of the content of the scientific articles published by these expert communities. One potential conclusion may be that less publication in dermatology suggests failure to adopt. However, this is not the case. More fine-grained analysis of journal article titles and abstracts showed that while the community of dermatologists broadly accepted the validity of TPMT test, this test is still highly contested in the communities of rheumatologists and, especially, gastroenterologists. This more intense debate that is captured by the cognitive mapping approach is also testified by the strongly worded titles of scientific articles and presence of a large number of author response/reply or comment articles to other articles that these two communities published.

A similar cognitive perspective can be built by using a map of which nodes represent academic

---

17 The visualisations of cognitive maps were produced by using VOSviewer 1.5.4 (van Eck and Waltman 2010).
Figure 6: Overlay of the publishing activity across the scientific disciplines. *Source: authors’ elaboration.*
journals (Leydesdorff et al., 2013). The map is specifically composed by 10,330 journals (nodes) — the different colours of nodes represent different clusters of journals, i.e. groups of journals of which the cross-citation patterns are similar. Figure 7, for example, illustrates the rapid diffusion of RNAi across this map. RNAi has specifically started to appear in journals in the basic biomedical science and subsequently it has diffused among a variety of discipline-specific journals. The Rao-Stirling diversity index (Stirling, 2007), measured on the set of journals of the map, provides further evidence of this rapid diffusion, especially when the index is compared with the other two emerging technologies on which we focused our analysis.

![Journal map of the RNAi case-study and diversity indexes for the considered three case-studies. Source: authors' elaboration.](image)

Perspectives on the cognitive dynamics can be also built by using MeSH terms — terms used to characterise the content of scientific articles in life science. These terms are assigned to articles in MEDLINE/PubMed through an intensive indexing process that is performed by examiners at the National Institute of Health (NIH). The terms are organised in a 16-branch tree, which can reach up to 12 levels of depth. Drawing from this classification, Leydesdorff et al. (2012) developed a MeSH map by using three branches: ”Diseases”, ”Chemicals and Drugs”, and ”Analytical, Diagnostics and Therapeutic Techniques and Equipment”. These can be interpreted as 'supply-side', 'demand-side' and 'infrastructure', respectively. The first two
levels of the tree were selected to build the basemap. The map is specifically composed by 822 MeSH terms (nodes) of which linkages reflect the (cosine) similarity according to co-occurrence of these terms in scientific articles. Each branch is marked on the map with a different colour: "Disease" is coloured red, "Chemicals and Drugs" green, and "Analytical, Diagnostics and Therapeutic Techniques and Equipment" blue (see Figure 8). Similarly to previous approaches, the publishing activity characterising a given emerging technology can be projected on this map to trace dynamics across three branches of the MeSH tree.

This approach, applied to the three case-studies, revealed different dynamics. RNAi, in line with previous results, 'globalises' across the set of the MeSH terms thus affecting many areas of the represented branches. On the contrary, HPV testing technology diffuses from the "Diseases" branch, specifically from "Tumor Virus Infections", into the "Analytical, Diagnostics and Therapeutic Techniques and Equipment" branch and eventually across the "Chemicals and Drugs" area. Yet, interestingly, in the last time window (2007-2011 period) scientific articles on HPV testing technology concentrate in the techniques and equipment area. This may reflect the efforts in developing competing HPV testing technologies. Results also show the specialisation of the TPMT case-study in more limited areas of the map, reflecting its narrow range of application (namely a handful of clinical niches).

Tracing the patenting activity of emerging technologies provides additional perspectives on the cognitive dynamics of the emergence process given the diverse incentives featuring in the creation process of scientific articles and patents. Scholars have developed techniques also to trace the dynamics of the patenting activities (e.g. Kay et al., 2014; Schoen et al., 2012). The nodes of these maps are technological classes that, as in the case of previous maps, are linked by cross-citation (cosine) similarity (Leydesdorff et al., 2015, 2014).

Figure 9, for example, depicts the overlays of RNAi patenting activity on the patent map based on technological areas as defined by the International Patent Classification (IPC). One can trace the dynamics in this space by moving across different levels of the classification (e.g. 3-digit, 4-digit). The patent map visualisation revealed the patenting activity of RNAi focused in specific areas of the technological space as biochemistry, organic chemistry, and medical science. While one would expect to observe intense patenting in these technological areas given the nature of RNAi, this also reveals how one may need to increase, from case to case, the granularity of the mapping (i.e. lower levels of IPC classes) to build informative and interpretative perspectives on the observed emerging technology.
Figure 8: Overlay of the publishing activity across the map of MeSH terms. *Source: authors’ elaboration.*
Mapping the emergence in the cognitive space may therefore reveal a number of relevant dynamics. These include the directions of diffusion of the given emerging technology across the key knowledge areas involved in emergence, how those areas may integrate or misalign, in which domains actors’ knowledge production processes are positioned, or, for medical innovation, what diseases a technology is addressing and by using what type of techniques and chemicals.

5 Discussion

In cases of the emergence of novel technologies, all actors and particularly policy makers have very incomplete knowledge, not knowing, for example, where the boundaries of the technology are, in which direction it is moving, and how it could move forward. In the face of uncertainty, ambiguity, or ignorance (Stirling, 2007), the emergence process should be investigated and analysed with strategic intelligence tools in order to support a more informed policy-making process. In this article, we have presented a set of overlay mapping approaches, applied those on three case-studies of emerging technologies in the biomedical domain, and shown the variety of 'intelligence' inputs for the policy-making process they can generate.

One of the advantages of the techniques we present is they facilitate analysis of the multifaceted, complex process of emergence across the geographical, social, and cognitive spaces (and combinations of these) rather than relying on a single perspective which may not reveal key trends. Overlay mapping can be therefore conceived as a monitoring system for emerging technologies that by tracking emergence and synthesising the complex information in relatively accessible visualisations can potentially favour the development of policy instruments of a higher scope, speed, accuracy, and reliability (Nightingale, 2003). Overlay data can be used for assessment of the dynamics of emerging technologies. The above discussed routines save overlay data (that, in most of the cases, are mathematically represented by vectors) into a relational database, which can be queried to generate statistics on the considered emerging technology.

However, the techniques’ efficacy and efficiency as a strategic intelligence tool are dependent on a set of choices that the analyst makes — although some choices may seem purely technical (e.g. level of aggregation), in practice they have an interpretative component that may also have important implications for the type of patterns one can observe (e.g. the granularity of results). Firstly, a key preliminary choice is to identify the boundary of emerging technologies, i.e. the delineation of the corpus. For analysts, deciding what to include and what to exclude
Figure 9: Patent map (IPC-based) for RNAi. Source: authors’ elaboration.
in the analysis is often a problematic exercise in the case of emergent sciences and technologies. In the case of RNAi, for example, one type of delineation portrayed this technology as already having reached a mature phase, whereas a broader definition suggested RNAi in a phase of rapid growth (Figure 2). To address this issue, the analyst should rely on interactions with experts in the field of the emerging technology under study.

Secondly, the analyst has to identify appropriate data sources. It is clear that using bibliographic data from publications and patents yields different types of information. Yet, those are only one form of research outputs. Many other possible dimensions such as products, services, changes in healthcare outcomes, are not easily accessible for this mapping process. This implies that some emerging technologies will be well represented in certain datasets and not others. This may also be the case for certain groups of actors — for example, academic organisations may publish more and patent less than private organisations, who may not have as high a propensity to publish papers as they do to apply for patents. Consideration of available datasets places a considerable onus on the analyst to find appropriate sources and not to over-interpret limited data from less appropriate sources. Studying biotechnological innovations in an IP dominated industry, where regulatory and peer community pressures ensure publication, is a considerable advantage from the point of view of data access and availability. However, other contexts may not be so well provisioned, with implications for the utility of the approaches discussed here. The extension of overlay mapping approach to other types of data sources represents an important challenge for future research especially when one considers the increasing attention towards the use of 'big data' (e.g. Thelwall et al., 2013). While data based on social media data may yield interesting information, these so-called ‘altmetrics’ techniques may be more sensitive to fads and hype than in the databases where the generation of outputs requires more effort (e.g. publications in WoS, MEDLINE/PubMed, and patent applications e.g. to USPTO).

Thirdly, the selection of the elements to be analysed from the records of the databases as well as the categories into which the elements are assigned may have a significant impact on the resulting analyses. From a patent record, for example, one can extract information about inventors, firms, technology classes, or location, which, as discussed, provide insights on the social, cognitive or geographical spaces of emergence, respectively. In the case of categories one needs to decide the level of aggregation (the granularity of the description) as well as the type of classification — i.e., whether a predetermined, top-down typology is used (e.g., the MeSH terms) or an emerging, bottom-up taxonomy.
Ideally, the choices of databases, elements, and categories can be informed by conceptual and theoretical frameworks. However, from a policy analyst’s standpoint, it may be also critical to monitor technological emergence even when there is lack of an explicit understanding of conceptual frameworks used — policy is needed under conditions of incomplete knowledge. Yet, even in these cases in which there is no explicit adoption of a conceptual framework, the choice of certain elements and categories in the analysis is privileging certain understanding over others, and the implicit assumptions that drive the analysis need to be made clear to ensure the maps are interpreted within an appropriate context. It is all too easy to misread a map.

For example, looking at collaborative networks of individual scientists implicitly places more attention on the social capital as a key factor in the emergence process (e.g. Nahapiet and Ghoshal 1998). If one looks at the disciplinary position of the technology it is likely to be assuming that integration of disparate knowledge is relevant as, for example, it has been perceived to be the case of nanotechnologies (e.g. Porter and Rafols 2009) and in RNAi (Leydesdorff and Rafols 2011), but not necessarily in HPV or TPMT testing technologies. In the absence of a clear understanding of the technology, one is advised to explore the phenomenon by using several perspectives, because one does not know in advance which one may turn out to be useful for understanding the relations of the emergence in place, or because the area in which the action is occurring are shifting over time.

Emerging technologies do not conform to established bodies of knowledge. They cut across pre-existing organisational and institutional units, challenging established managerial and policy practices. As a result, one key demand from analysts and decision-makers is a description of the types of interdisciplinarity or convergence evident in emergence (Schmidt 2007), often related with specific visions and expectations (Beckert et al. 2007; Roco and Bainbridge 2002). In this regard, building multiple perspectives on the process of emergence requires looking across databases where emerging technologies ‘tumble’ with different representations that stress different attributes — such as geographical addresses, WoS Categories, patent classes.

The databases are retained and organized in different contexts with relative institutional rigidities. For example, relating patents to publications in terms of ’non-patent literature references’ requires professional skills and cannot be done on a large scale without substantive investments. Thus, while the mapping and overlay techniques allow one to use the same or similar search strings across these databases more research is required to increase the integration of datasets. This can already be seen in attempts by US funders, such as the National Institutes
of Health (NIH), to link research grants, patents and publications.\(^\text{18}\)

In summary, while overlay maps are a potentially useful tool for informing decision-making, the adequacy of this tool for mapping emerging technologies is quite sensitive to a variety of choices (such as boundary of technology, data sources used, elements and level of granularity of analyses). Because the maps are generated during emergence, there is a high degree of uncertainty and ambiguity regarding the choices to be made. As a result, these maps are not expected to accurate or even very reliable – they are just heuristic pictures of shapes yet in the making. The purpose of using multiple dimensions to investigate emergence is to help in spotting robust insights against patterns that are the result of analytical artifacts.

6 Conclusions

This articles shows how overlay mapping can function as a tool for strategic intelligence in the context of policy processes on the governance of emerging technologies. These scientometric techniques are becoming increasingly user-friendly and accessible, allowing outputs to be produced using recent or historical trend data in order to inform perspectives on emergence with relative low efforts of analysts and decision-makers. Relevant dynamics are illustrated across the geographical, social, and cognitive spaces, i.e. in terms of geographical distribution of knowledge processes, collaborative interactions among geographical areas and organisational actors, and scientific and technological domains emerging or involved. Overlay mapping can also crosscut sources of data and move across units of analysis (organisations, cities, disciplines, technologies and sub-fields). This flexibility and diverse granularity favour the comparison of the results from different policy-making contexts.

We illustrated our argument conducting three case-studies of emerging technologies in the biomedical domain. The case-studies, being heterogeneous in terms of scale, applications, and phase of development, allowed us to explore the use of these techniques in distinct circumstances. Results showed how the overlay mapping approach provides both overview and the possibility to focus by zooming in on specific developments and phenomena. These maps are ‘interactive’ and allow the user to upload their own data of interest.

In the case of HPV and TPMT testing technologies, we showed how different facets and dynamics of the emergence process can be revealed by moving across the geographical and

\(^{18}\)See, for instance, [http://projectreporter.nih.gov/reporter.cfm](http://projectreporter.nih.gov/reporter.cfm)
social spaces from the city-level to organisation-level units of analysis. The maps also enable the user to compare different technologies and alternatives in terms of strengths and weaknesses at the portfolio level. As in the case of RNAi, the mapping can reveal cases of emerging technologies initially concentrated at a few places with major players hosting the sets of relevant capabilities, competences, and network relationships, and then diffusing across a number of actors according to preferential attachment mechanisms (Leydesdorff and Rafols 2011). On this ‘journey’ of the emerging technology, changes in the involved domains of science and technology — especially in the case of RNAi, the diffusion of laboratory research tools in other specialties and disciplines — can also be revealed with the cognitive mapping.

Overlay maps has some potential problems that need to be addressed during the study, namely: the delineation issues associated with emerging technologies, the need of some qualitative background on the studied technology for the interpretation and refinement of the mapping, the limitations of the data that can be used (mainly publications and patents). If these problems prove manageable, overlay mapping can function as a heuristic tool that can reveal key trends and allow stakeholders to bring relevant data to theoretical and policy debates. Maps throw new light on relationships and raise additional questions for analysts and decision-makers to address, thus opening up the space of discussion. The resulting perspectives on emergence, according to the set of choices an analyst takes, may help debate in a timely manner the directions of further investigation as well as feeding into political discourse about stumbling blocks ahead and possible openings in the landscapes.

Acknowledgements

The authors acknowledge the support of UK Economic and Social Research Council (award RES-360-25-0076 - "Mapping the Dynamics of Emergent Technologies") in the development of the case-studies. The authors are also grateful to the US National Science Foundation (award #1064146 - "Revealing Innovation Pathways: Hybrid Science Maps for Technology Assessment and Foresight") for the support in the development of the mapping techniques. The findings and observations contained in this paper are those of the authors and do not necessarily reflect the funders’ views. The authors are grateful to Paul Nightingale, Stefan Kuhlmann, and Peter Stegmaier for their feedback on earlier versions of this paper. Previous versions of this paper were presented at the 2013 ISSI Conference, 2013 EU-SPRI Conference, 2013 Atlanta Conference,
2013 DRUID Conference, and "2013 Data Science, Tech Mining and Innovation" workshop in Manchester.

References

Alexander, J., Chase, J., Newman, N., Porter, A., and Roessner, J. (2012). Emergence as a conceptual framework for understanding scientific and technological progress. In 2012 Proceedings of PICMET’12: Technology Management for Emerging Technologies, pages 1286–1292.

Beckert, B., Blümel, C., and Friedewald, M. (2007). Visions and realities in converging technologies. Exploring the technology base for convergence. Innovation: The European Journal of Social Science Research, 20(4):375–194.

Börner, K. (2010). Atlas of Science. Vizualizing What We Know. MIT Press, Cambridge, MA.

Bornmann, L. and Leydesdorff, L. (2011). Which cities produce more excellent papers than can be expected? A new mapping approach, using Google Maps, based on statistical significance testing. Journal of the American Society for Information Science and Technology, 62(10):1954–1962.

Bornmann, L. and Waltman, L. (2011). The detection of "hot regions" in the geography of science. A visualization approach by using density maps. Journal of Informetrics, 5(4):547–553.

Callon, M., Courtial, J. J. P., Turner, W. A., and Bauin, S. (1983). From translations to problematic networks: An introduction to coword analysis. Social Science Information, 22(2):191–235.

Collingridge, D. (1980). The Social Control of Technology. Open University Press, Milton Keynes.

Crane, D. (1972). Invisible Colleges: Diffusion of Knowledge in Scientific Communities. The University of Chicago Press, Chicago.

David, P. (1985). Clio and the economics of QWERTY. The American Economic Review, 75(2):332–337.

Day, G. S. and Schoemaker, P. J. H. (2000). Avoiding the pitfalls of emerging technologies. California Management Review, 42(2):8–33.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature, 391(6669):806–811.

Garfield, E. (1979). Citation Indexing: Its Theory and Application in Science, Technology, and Humanities. John Wiley, New York.

Garfield, E., Sher, I. H., and Torpie, R. J. (1964). The Use of Citation Data in Writing the History of Science. Institute for Scientific Information, Philadelphia.

Gibbert, M., Ruigrok, W., and Wicki, B. (2008). What passes as a rigorous case study? Strategic Management Journal, 29(13):1465–1474.

Giddens, A. (1984). The Constitution of Society. University of California Press, Berkeley, CA.
Glänzel, W. and Schubert, A. (2004). Analysing scientific networks through co-authorship. In Moed, H., Glänzel, W., and Schmoch, U., editors, Handbook of Quantitative Science and Technology Research, pages 257–276. Kluwer, Dordrecht.

Glänzel, W. and Thijs, B. (2011). Using core documents for detecting and labelling new emerging topics. Scientometrics, 91(2):399–416.

Griliches, Z. (1994). Productivity, R&D, and the data constraint. The American Economic Review, 84(1):1–23.

Haussecker, D. (2008). The business of RNAi therapeutics. Human Gene Therapy, 19(5):451–462.

Haussecker, D. (2012). The business of RNAi therapeutics in 2012. Molecular Therapy Nucleic Acids, 1(2):e8.

Hogarth, S., Hopkins, M. M., and Rodriguez, V. (2012). A molecular monopoly? HPV testing, the Pap smear and the molecularisation of cervical cancer screening in the USA. Sociology of Health & Illness, 34(2):234–250.

Hopkins, M. M., Ibarreta, D., Gaiser, S., Enzing, C. M., Ryan, J., Martin, P. A., Lewis, G., Detmar, S., Marle, M. E. v. d. A.-v., Hedgecoe, A. M., Nightingale, P., Dreiling, M., Hartig, K. J., Vullings, W., and Forde, T. (2006). Putting pharmacogenetics into practice. Nature Biotechnology, 24(4):403–410.

Hu, G., Carley, S., and Tang, L. (2012). Visualizing nanotechnology research in Canada: evidence from publication activities, 1990-2009. The Journal of Technology Transfer, 37(4):550–562.

Hung, S.-C. and Chu, Y.-Y. (2006). Stimulating new industries from emerging technologies: challenges for the public sector. Technovation, 26(1):104–110.

Hunt, D., Nguyen, L., and Rodgers, M. (2007). Patent Searching: Tools and Techniques. John Wiley & Sons, Hoboken, New Jersey.

Irvine, J. and Martin, B. R. (1985). Basic research in the East and West: A comparison of the scientific performance of high-energy physics accelerators. Social Studies of Science, 15(2):293–341.

Jaffe, A. B. and Trajtenberg, M. (2002). Patents, Citations & Innovations: A Window on the Knowledge Economy. MIT Press, Cambridge, MA.

Joerges, B. and Shinn, T. (2002). Instrumentation Between Science, State and Industry. Kluwer Academic Publishers, Dordrecht.

Kay, L., Newman, N., Youtie, J., Porter, A. L., and Rafols, I. (2014). Patent overlay mapping: Visualizing technological distance. Journal of the Association for Information Science and Technology, 65(12):2432–2443.

Kessler, M. M. M. (1963). Bibliographic coupling between scientific papers. American Documentation, 14(1):10–25.

Klavans, R. and Boyack, K. W. (2009). Toward a consensus map of science. Journal of the American Society for Information Science and Technology, 60(3):455–476.

Kooiman, J. (1993). Modern Governance. New Government-Society Interactions. Sage, London.
Kuhlmann, S., Boekholt, P., Georgiou, L., Guy, K., Heraud, J.-A. A., Laredo, P., Lemola, T., Loveridge, D., Luukkonen, T., Moniz, A., Polt, W., Rip, A., Sanz-Menendez, L., and Smits, R. (1999). Improving distributed intelligence in complex innovation systems. MPRA Working Paper.

Latour, B. (1993). The Pasteurization of France. Harvard University Press, Cambridge, MA.

Laudel, G. (2002). What do we measure by co-authorships? Research Evaluation, 11(1):3–15.

Leydesdorff, L. (1989). Words and co-words as indicators of intellectual organization. Research Policy, 18(4):209–223.

Leydesdorff, L., Alkemade, F., Heimeriks, G., and Hoekstra, R. (2015). Patents as instruments for exploring innovation dynamics: geographic and technological perspectives on photovoltaic cells. Scientometrics, 102(1):629–651.

Leydesdorff, L. and Bornmann, L. (2012). Mapping (USPTO) patent data using overlays to Google Maps. Journal of the American Society for Information Science and Technology, 63(7):1442–1458.

Leydesdorff, L., Kushnir, D., and Rafols, I. (2014). Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC). Scientometrics, 98(3):1583–1599.

Leydesdorff, L. and Persson, O. (2010). Mapping the geography of science: Distribution patterns and networks of relations among cities and institutes. Journal of the American Society for Information Science and Technology, 61(8):1622–1634.

Leydesdorff, L. and Rafols, I. (2011). The local emergence and global diffusion of research technologies: An exploration of patterns of network formation. Journal of the American Society for Information Science and Technology, 62(5):846–860.

Leydesdorff, L., Rafols, I., and Chaomei, C. (2013). Interactive overlays of journals and the measurement of interdisciplinarity on the basis of aggregated journal-journal citations. Journal of the American Society for Information Science and Technology, 64(12):2573–2586.

Leydesdorff, L., Rotolo, D., and Rafols, I. (2012). Bibliometric perspectives on medical innovation using the medical subject Headings of PubMed. Journal of the American Society for Information Science and Technology, 63(11):2239–2253.

Lundin, P. (2011). Is silence still golden? Mapping the RNAi patent landscape. Nature Biotechnology, 29(6):493–497.

Martin, B. R. (1995). Foresight in science and technology. Technology Analysis & Strategic Management, 7(2):139–168.

Nahapiet, J. and Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. Academy of Management Review, 23(2):242–266.

Narin, F. (1976). Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity. National Science Foundation, Washington, DC.

Nightingale, P. (2003). Capacity utilization revisited: software, control and the growth of large technical systems. Industrial and Corporate Change, 12(3):477–517.

Noyons, C. (2004). Science maps within a science policy context. In Moed, H., Glänzel, W., and Schmoch, U., editors, Handbook of Quantitative Science and Technology Research, pages 237–255. Kluwer, Dordrecht.
Noyons, E. (2001). Bibliometric mapping of science in a science policy context. *Scientometrics*, 50(1):83–98.

Porter, A. L. and Detampel, M. J. (1995). Technology opportunities analysis. *Technological Forecasting & Social Change*, 49(3):237–255.

Porter, A. L. and Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, 81(3):719–745.

Porter, A. L., Roessner, J. D., Jin, X.-Y., and Newman, N. C. (2002). Measuring national emerging technology capabilities. *Science and Public Policy*, 29(3):189–200.

Powell, W. W. (1990). The Transformation of Organizational Forms: How Useful is Organisation Thoery in Accounting for Social Change. In Friedland, R. O. and Robertson, A., editors, *Beyond the Marketplace: Rethinking Economy and Society*. Walter de Gruyter, New York, NY, USA.

Price, D. (1965). Is technology historically independent of science? A study in statistical historiography. *Technology and Culture*, 6(4):553–568.

Rafols, I., Porter, A. L., and Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9):1871–1887.

Rheinberger, H.-J. (1997). *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford University Press, Palo Alto, CA.

Rip, A. and Van Amerom, M. (2010). Emerging De Facto Agendas Surrounding Nanotechnology: Two Cases Full of Contingencies, Lock-outs, and Lock-ins. In Kaiser, M., Kurath, M., Maasen, S., and Rehmann-Sutter, C., editors, *Governing Future Technologies*, volume 27 of *Sociology of the Sciences Yearbook*, pages 131–155. Springer Netherlands, Dordrecht.

Robinson, D. K. R., Ruivenkamp, M., and Rip, A. (2007). Tracking the evolution of new and emerging S&T via statement-linkages: Vision assessment in molecular machines. *Scientometrics*, 70(3):831–858.

Roco, M. and Bainbridge, W. (2002). *Converging technologies for improving human performance. Nanotechnology, biotechnology, information technology and cognitive science*. National Science Foundation, Arlington, VA, US.

Rotolo, D., Hicks, D., and Martin, B. (2015). What is an emerging technology? *SPRU Working Paper Series*, in press.

Schmidt, J. (2007). Knowledge politics of interdisciplinarity. Specifying the type of interdisciplinarity in the NSF’s NBIC scenario. *Innovation: The European Journal of Social Science Research*, 20(4):313–328.

Schoen, A., Villard, L., Laurens, P., Cointet, J.-P., Heimeriks, G., and Alkemade, F. (2012). The network structure of technological developments: Technological distance as a walk on the technology map. In *2012 Science & Technology Indicators (STI) Conference*, Montreal.

Small, H. (1973). Co-citation in the scientific literature: a new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4):265–269.

Small, H. (1977). A co-citation model of a scientific specialty: A longitudinal study of collagen research. *Social Studies of Science*, 7(2):139–166.
Small, H., Boyack, K. W., and Klavans, R. (2014). Identifying emerging topics in science and technology. *Research Policy*, 43(8):1450–1467.

Small, H. and Garfield, E. (1985). The geography of science: disciplinary and national mappings. *Journal of Information Science*, 11(4):147–159.

Stirling, A. (2007). A General framework for analysing diversity in science, technology and society. *Journal of The Royal Society Interface*, 4(15):707–719.

Sung, J. J. and Hopkins, M. M. (2006). Towards a method for evaluating technological expectations: Revealing uncertainty in gene silencing technology discourse. *Technology Analysis & Strategic Management*, 18(3-4):345–359.

Thelwall, M., Haustein, S., Larivière, V., and Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PloS one*, 8(5):e64841.

van Eck, N. J. and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2):523–538.

Van Lente, H. and Rip, A. (1998). The rise of membrane technology. From rethorics to social reality. *Social Studies of Science*, 28(2):221–254.

Verbong, G. and Geels, F. (2007). The ongoing energy transition: Lessons from a socio-technical, multi-level analysis of the Dutch electricity system (1960-2004). *Energy Policy*, 35(2):1025–1037.

Wagner, C. S. (2008). *The New Invisible College*. Brookings Institution Press, Washington, DC.

Waltman, L. and van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal American Society for Information Science and Technology*, 63(12):2378–2392.

Wassermann, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, MA.

Yin, R. K. (2003). *Case Study Research: Design and Methods (Fourth Edition)*. Sage Publications, Thousand Oaks, US, third edition.

Zitt, M. and Bassecoulard, E. (2006). Delineating complex scientific fields by an hybrid lexical-citation method: An application to nanosciences. *Information Processing & Management*, 42(6):1513–1531.

zur Hausen, H. (1987). Papillomaviruses in human cancer. *Applied Pathology*, 5:19–24.