Chapter 10
Visual Analysis of Patent Data Through Global Maps and Overlays

Luciano Kay, Alan L. Porter, Jan Youtie, Nils Newman, and Ismael Ràfols

Abstract Visual analytics has been increasingly used to help to better grasp the complexity and evolution of scientific and technological activities over time, across science and technological areas and in organisations. This chapter presents general insights into some important fields of expertise such as mapping, network analysis and visual analytics applied to patent information retrieval and analysis. We also present a new global patent map and overlay technique and illustrative examples of its application. The concluding remarks offer considerations for future patent analysis and visualisation.

10.1 Introduction

Visual analytics has been increasingly used to help to better grasp the complexity and evolution of scientific and technological activities over time, across science and technological areas and in organisations. New and diverse analysis

L. Kay (✉)
Center for Nanotechnology in Society (CNS) – ISBER, University of California Santa Barbara, 2201 North Hall, Santa Barbara, CA, 93106-2150, USA
e-mail: luciano.kay@ucsb.edu

A.L. Porter
School of Public Policy, Georgia Institute of Technology, Atlanta GA & Search Technologies, Norcross, GA, USA
e-mail: alan.porter@isye.gatech.edu

J. Youtie
Enterprise Innovation Institute & School of Public Policy, Atlanta, GA, USA
e-mail: jan.youtie@innovate.gatech.edu

N. Newman
Intelligent Information Services Corporation, Atlanta, GA, USA
e-mail: newman@iisco.com

I. Ràfols
Ingenio (CSIC-UPV), Universitat Politècnica de València, València, Spain

SPRU, University of Sussex, Brighton, UK
e-mail: i.rafols@ingenio.upv.es

M. Lupu et al. (eds.), *Current Challenges in Patent Information Retrieval*,
The Information Retrieval Series 37, DOI 10.1007/978-3-662-53817-3_10

and mapping methods, increasing computing power and new software and layout algorithms enable this and support patent analysis aimed at understanding a range of innovation-related phenomena.

This chapter presents general insights into some important fields of expertise such as mapping, network analysis and visual analytics applied to patent information retrieval and analysis. We discuss broader aspects and issues of patent mapping, including the development of global maps and overlays in the context of information retrieval, exploration and analysis for patent corpora, and their similarities/differences to similar approaches applied to scientific literature. The need for development of tools to benchmark and capture temporal change of organisational innovation activities, or patterns of technological change, also motivates this work.

We also present the new global patent map and overlay technique we recently developed [1]. Our visualisation approach is a logical extension of experience acquired with science overlay maps [2] and opens up new avenues for understanding patent landscapes, which as we will see markedly differ from scientific landscapes. To illustrate the kind of analytical support offered by this approach, we discuss the core structure of our global patent map and apply patent overlay maps to benchmark the nanotechnology-related patenting activities of companies. We conclude this chapter by offering some remarks and considerations for future patent analysis and visualisation.

10.2 Patent Information Retrieval and Analysis

Patent analysis plays key roles in competitive technical intelligence (CTI) [3]. The multipart technological innovation CTI 'puzzle' comprises both empirical information and expert analysis to inform empirical search, refining and interpretation. Patents provide an important piece of empirical information in the form of compilations of large numbers of records for 'landscaping', i.e. a macro perspective—our focus here—as well as in-depth treatment of a small number of patent documents for micro-perspective analyses [4]. Other complementary empirical information comprise research publication search compilations/reviews and roadmaps and business-related content (e.g. trade publications, policy documents, popular press treatment).

One implication of such an 'innovation systems' perspective [5, 6] is to see potential value in ways to combine multiple information types. We generally seek innovation indicators [7]. Patent mapping provides a visual component to enrich various innovation system analyses [8, 9]. In particular, we think patent mapping can complement science overlay mapping to enrich understanding of research and development (R&D) activities, particularly for engagement of subareas and maturation patterns [10]. Leydesdorff et al. [11] have devised patent overlay maps and Leydesdorff et al. [12] illustrate their potential in exploring innovation dynamics in areas such as photovoltaic technologies, both over geographical regions and over topical regions.

The macro-patent analyses that patent overlay mapping serves seek to discern patent activity patterns with implications for innovation. These can inform corporate investment decisions via intelligence about key competitors' perceived trajectories. Chen [13], for example, shows patent 'landscape' maps created by Boyack [14] changing over time. These aid CTI in tracking competitor interest evolution. Alternatively, analyses can contribute to policy discourse by profiling national positions and potential. Of course, patents are an imperfect lens on innovation—they reflect invention, and that unevenly, as patent practices vary greatly by industrial sector and country (cf. [15]).

The unit of analysis is a collection of patent information relating to a target topic. That collection typically contains patent abstract records, not full-text patent documents. Those would be gathered via a search strategy applied to one or more databases (e.g. *Derwent World Patent Index*, *EPO PATSTAT*). The use of diverse text data sources (sometimes with varying language usage, technical terms or machine-translated documents) implies that a well-crafted search is essential, and it must address one or more research questions. In the private sector CTI realm, those questions tend to focus on either a key competitor or a few of them and their intents regarding a certain technology or application area. In public sector or academic treatments, focus is more apt to be on an emerging technology, cast broadly.

The search strategy can be conducted using Boolean term searching (combining key terms, often delimited by proximity conditions). Or, the search can rely on patent classification specifications such as International Patent Classifications (IPCs). Often a combination search query is most effective. Search quality is essential and criteria centre on how best to address the driving research questions. In general, macro-scale profiling leans towards inclusive search, thereby providing the option of further analyses by refining to subsets of the data retrieved.

Once patent abstract data are retrieved, the analyst faces notable challenges in 'getting the data right'. We have found it fruitful to engage domain experts to review initial search set patterns, particularly top terms and phrases, to spot flaws or gaps, and suggest ways to improve our search queries. For instance, in recent work on nano-enabled drug delivery (NEDD) we removed some 5 % of the search set concerning agriculture [16]. The next stage entailed data cleaning. This can vary enormously in scale of effort, contingent on the sensitivities in addressing the driving analytical questions. For broad patent landscaping, we want to get a representative sample. For massive searches, this implies tradeoffs in scope—e.g. maybe reducing the search time frame. For the NEDD analysis, with the purpose of visualising patenting distribution over patent categories, we set aside search terms concerning cancer to reduce distortion in not specifying other target diseases.

Analysts can use diverse forms of patent analysis with a varying degree of complexity. These range from the generation of lists of patent records and co-occurrence matrices (lately made simpler thanks to text mining software) to more complex clustering and mapping of patent data. Lists can filter patent records by given criteria or fields, and matrices help to find relationships resulting from, for example, co-occurrence of keywords in patent titles and abstracts. Document clustering allows identifying topics in patent literature and patent mapping provides windows on the pattern of invention. Geo-mapping of inventors and/or patent assignees can illuminate areas of strength (e.g. for national comparisons). Geo-mapping of patent authority activity, particularly when staged over time periods, may elucidate relative market potential. And, patent overlay mapping applications as presented here contribute insight into component technologies, as well as market sectors being engaged.

Citation analyses deserve mention as well. One may gain useful technology transfer insights by considering the patents (and/or literature) cited by a target patent set and the patents that cite such a target set. The latter are especially affected by patent time lags. With regard to the patent mapping we present later in this chapter, one needs to know the IPC of the cited (or citing) patents. That requires additional layers of search and retrieval.

10.3 Visual Analytics and Overlays

The visualisation of knowledge or technological landscapes has been a prominent part of publication and patent analyses since their origins [17, 18]. Only in the last decades, however, improvements in computational power and algorithms have allowed the creation of large maps covering a full database, the so-called global maps of science and technology (see overviews by [2, 19]). This in part has led to a proliferation of global maps ([20–25]; e.g. see [26, 27]).

Science maps or *scientograms* are the visualisation of the relations among areas of science using network analysis algorithms. Visualisation procedures for science maps have generally been used to explore and visually identify scientific frontiers, grasp the extent and evolution of scientific domains and analyse the frontiers of scientific research change [28]. Science mapping efforts have also been used to inspire cross-disciplinary discussion to find ways to communicate scientific progress.[1]

A patent map, on the other hand, is a symbolic representation of technological fields that are associated with relevant themes. Technological fields are positioned in the map so that similar fields are situated nearby and dissimilar components are situated at a distance. Their construction uses similar algorithms to those used to visualise the relations among scientific disciplines. Patent maps help to explore and visually identify areas of technology development concentration, and they can illuminate increasing or diminishing patenting activity over time. In this way, patent maps can inform R&D management, competitive intelligence and policy decision-making. A key characteristic of patent maps is the ability to graphically represent ‘technological distance’ or the extent to which a set of patents reflects different types of technologies [29]. Technological distance, often proxied by patent

[1] See, for example, the Mapping Science website at http://www.scimaps.org/

categories, with patents in a given patent category being considered more similar to one another than to those in other patent categories [30, 31], provides a measure of interrelatedness and potential innovation opportunities.

Science and technology maps complement other methodological approaches to data analysis. They can help to interpret and find meaning in complex data by transforming abstract and intangible datasets into something visible and concrete [13]. Scholars have pursued diverse approaches to scientific publication and patent record-level analysis to create global maps of science and technology. These serve to characterise the proximity and dependency of scientific areas (e.g. [19, 32]) and technological areas (e.g. preliminary work in Boyack and Klavans [33] and related approaches by Schoen et al. [9] and Leydesdorff et al. [11]. Notwithstanding the range of classification and visualisation algorithms, the resulting global maps have been generally 'stable', at least in terms of their main disciplinary or technological areas and their relationships. Still, the comparison of results of diverse approaches is important to test the robustness of patterns observed. Without significant consensus on the shape and relative position of science and technological categories, global maps are meaningless as stable landscapes needed to compare, for example, organisational or technological subsets.

The relative structural stability of global maps suggested their use as a base map over which to compare the technological distribution of specific organisations, in the same way that we may compare the distribution of different plant species or multinationals over the world map. This led to the idea of 'overlays' (or a process of 'layering' or 'stacking') by which global maps can be combined with additional layers that visually represent subsets of scientific publication and patent data. Overlays help to understand the particular scientific and technological thrusts and areas of concentration of R&D actors [2]. For example, a company's patent portfolio can be 'overlaid' on the base map. This process provides a visual tool to interpret the multidimensional relationships among the patent categories in the company's patent portfolio.

10.4 Visualising Innovation Pathways and Technology Development Concentrations

Our research has recently involved the creation of a new global patent map and overlay technique [1]. Our patent map is constructed from a similarity matrix based on citing-to-cited patents—i.e. a matrix that reflects similarities among IPC categories in how patents cite each other. The similarity measures are calculated from correlation functions among fields according to citations among patent categories. This multidimensional matrix is projected onto a two-dimensional space, which becomes our 'base map'. A user can then 'overlay' subsets of patent data—representing different types of technological fields, organisations or geographical regions—on top of the base map to understand the particular technological thrusts and areas of concentration of these entities.

While there have been other patent maps that use IPC categories (e.g. see [12, 17]), they share two main weaknesses, which our approach addresses. The first is the reliance on analysing patents at a given IPC level; 3-digit (class) and 4-digit (subclass) are the most commonly used levels. Patents are not equally distributed across IPC three-digit or IPC four-digit categories, however, so one experiences the problem of not being able to distinguish fields in classes that attract a huge number of patents—such as Medical or Veterinary Sciences (A61)—by staying within the confines of the existing IPC administrative structure. The second is assuming that patents in a given section of the IPC system are alike. For example, even though Medical or Veterinary Sciences (A61) and Hats (A42B) are both in Section A 'Human Necessities', they are not really that similar. 'Medical or Veterinary Sciences' is actually more similar to Organic Chemistry (C07), even though Organic Chemistry falls in Section C 'Chemistry, Metallurgy'. The approach presented in this chapter compensates for these issues by (1) disaggregating IPC categories and (2) reforming them based on citing-to-cited reference patterns. In addition, we remove some patent categories with fewer than 1000 patents to enable better ability to distinguish patterns in those categories with a higher propensity for patenting.

Our global patent map is based on citing-to-cited relationships among IPCs of European Patent Office (EPO) patents from 2000 to 2006. This period was chosen because of its stability with respect to IPC 7 categories. IPC 7, at the time we conducted this study, represented the longest period of stable classification. Future work would involve comparing patent overlay maps based on IPC 7 and future classification systems such as IPC 8 or Cooperative Patent Classification (CPC) systems, but first, the project team needed to make sure it could produce a mapping process with a stable set of categories. The dataset containing IPC relationships, extracted from the PATSTAT database version available in the fall of 2010, represents more than 760,000 patent records in more than 400 IPC categories. This data range begins with patent EP0968708 (which was published in January 2000) and ends with patent EP1737233 (published in December 2006).

A key part of our methodology involves disaggregating, then folding IPC categories up into the next highest level of aggregation to create relatively similar sized categories. This solution comprises three rules:

1. For IPC categories with large population, use the smallest subgroup level.
2. For small population IPC categories, aggregate up to general group level, subclass or class.
3. Establish a floor cut-off and drop very small aggregated populations.

As a result, IPC categories with instance counts greater than 1000 in the data set were kept in their original state. Those categories with instance counts less than 1000 were folded up to the next highest level until the count exceeded 1000 or the class level was reached. During the folding, any other IPC categories with counts exceeding 1000 in the same branch were left out of the folding count. If at the class

Table 10.1 Data pre-processing to group IPC categories, selected examples[a]

Original IPC in data set	Catchwords	Original record count
A61B	Diagnosis, surgery, identification	25,808
Authors' process splits this out into:		
A61B 5/00	Measuring for diagnostic purposes	1415
A61B 17/00	Surgical instruments, devices or methods, e.g. tourniquets	1493
A61B 19/00	Instruments, implements or accessories for surgery or diagnosis not covered by any of the groups	1444
And a remainder:		
A61B[b]		21,456

[a]Each IPC with an instance count greater than 1000 was kept in its original state
[b]Each IPC with an instance count less than 1000 was folded up to the next highest level until the count exceeded 1000 or the class level was reached

level (i.e. three-digit), the population was less than 1000, the IPC code was dropped for being too small to map. Table 10.1 illustrates this approach for the four-digit IPC class A61K.

This pre-processing yields IPC categories at the class (three-digit), subclass (four-digit), main group (five-digit) and subgroup (seven- and eight-digit) levels, with levels that ensure broadly similar numbers (i.e. within two orders of magnitude) of patents across categories. The next step involves extracting from PATSTAT the patents cited by the target records. The IPCs of those patents are mapped to the 466 IPC categories. Some of the patents cited by those in our IPC 7 data set were published under previous categorisation systems; however, this spillover does not lead to any problems from a categorisation standpoint because IPC integrates prior categorisations into more recent versions. The result of this data collection allows the creation of a table containing, in each row, sets of patent number, IPC number, cited patent number and cited IPC number.

The final data processing steps involve generating a cosine similarity matrix among citing IPC categories (using conventional cosine similarity normalised by the square root of the squared sum) and then factor analysis of the IPC categories (following the method used in global science maps by Leydesdorff and Rafols [24]. A factor analysis of the citing-to-cited matrix among IPC categories is then used to consolidate the 466 categories into 35 'macro-patent categories'. We tested different factor solutions ranging from 10 to 40 categories. The 35-factor solution had the greatest face validity, allowing a convenient classification of the IPC categories and easier interpretation. These 35 factors form the basis for colour-coding the 466 categories that are represented in visualisations. The visualisations also require converting IPC codes to succinct text labels, which we did by shortening lengthy IPC definitions. These IPC category labels were then used as a basis for creating descriptors for each factor as shown in the maps.

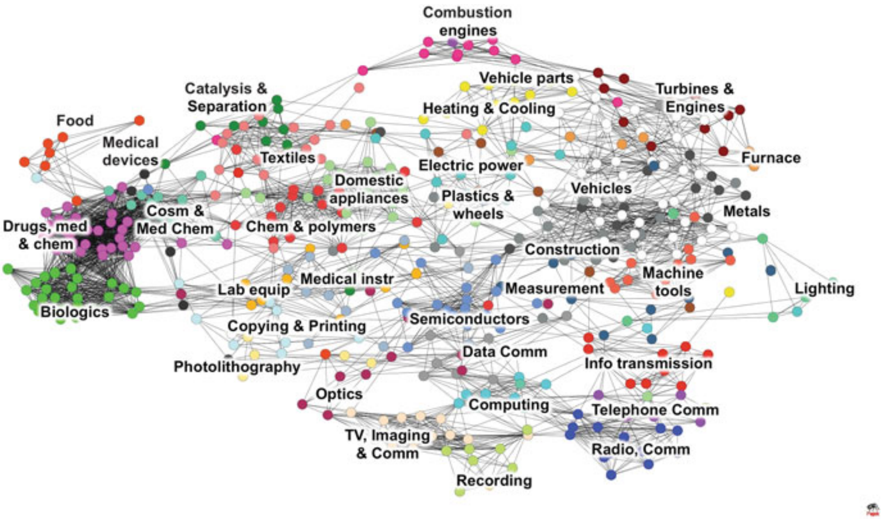

Fig. 10.1 Full patent map of 466 technological categories and 35 technological areas. Lines represent relationships between technological categories (the darker the line, the shorter the technological distance between categories)

The full map of patents shows all 466 categories in a Kamada-Kawai layout (using the software Pajek[2]) that represents technological distances and groups of technologies in each of the 35 factors or technological areas shown with the same colour (Fig. 10.1). Label and colour-related settings were adjusted to produce a reasonably clear map and facilitate its examination. The map suggests three broad dimensions of patenting interrelationships based on the overall position of technological areas. The left side of the map represents bio-related patents, including food, medicine and biology. The lower right part of the map includes semiconductor, electronics and information and communications technologies (ICT). The upper right portion of the map is primarily comprised of automotive and metal-mechanic-related technology groups.

To illustrate and test the application of patent map overlays, two corporate data sets of nanotechnology patent applications have been created for Samsung and DuPont, using data from the Georgia Tech Global Nanotechnology databases in the same time period (2000–2006). The visual examination of maps shows nanotechnology development foci that vary across companies (even for those in similar industry sectors) and different patenting activity levels for the studied period. The two overlays presented herein appear diversified and encompass a number of technological areas. The patent overlay created for Samsung, for example, shows activity concentrated on semiconductors and optics, with a notable level

[2]This software is free for non-commercial use: http://pajek.imfm.si

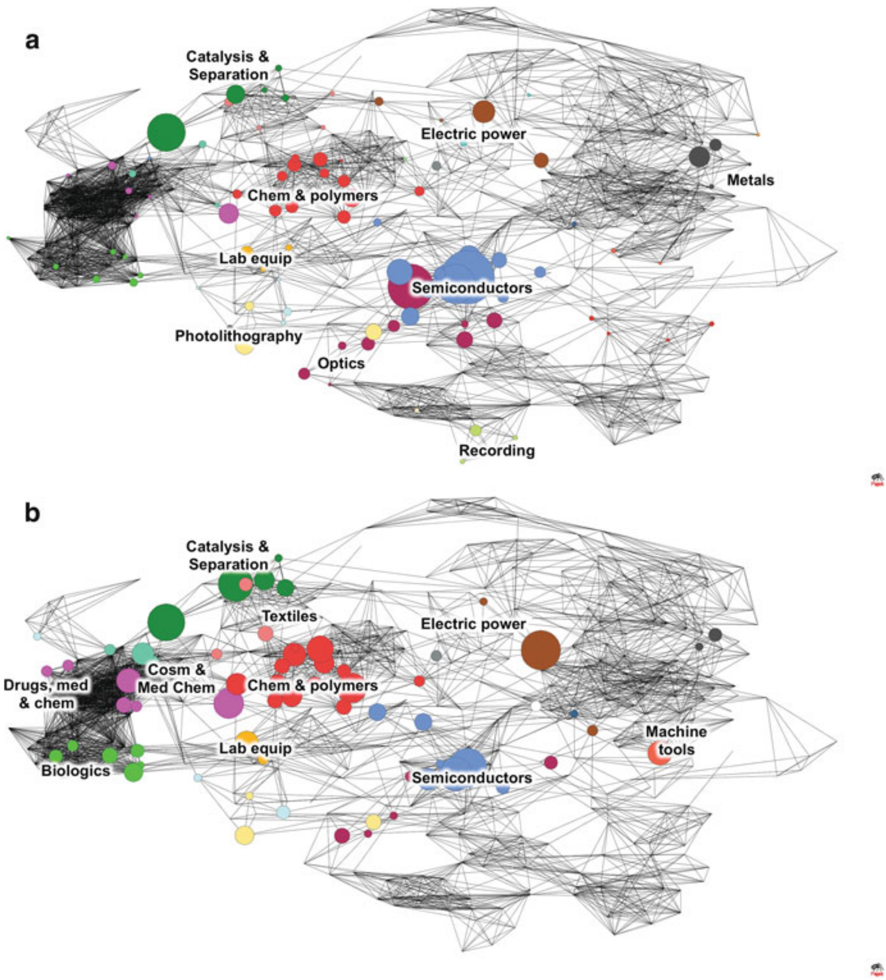

Fig. 10.2 Patent overlays applied to company benchmarking. (**a**) Samsung. (**b**) DuPont. The size of nodes is proportional to the number of patent applications in the corresponding technology group

of patenting activity across other areas as well (Fig. 10.2a). The company also has some prominent activity on technological areas broadly defined as Catalysis & Separation, Photolithography, and Chemistry & Polymers. DuPont, on the other hand, focuses on drugs, medicine and chemistry, chemistry and polymers as well as biologics (Fig. 10.2b). According to our overlays, this company has a portfolio of patent applications that is even more diversified, but it is also less active in terms of patenting activity, than Samsung.

10.5 Visualisations and Decision-Making Support

Visualisations can support R&D management, competitive intelligence and policy decision-making. Three main aspects of patent maps reveal the level of support they can offer: (1) map structures, (2) patterns of interconnection and (3) patent concentration.

The first aspect is connected with map structures, patent classifications and the challenge of relying on them for patent mapping. As technology changes, technology-oriented applications may draw from patents in different hierarchical categories and subsequently lead to further diversity in patents that cite patents in these categories. This requires making a distinction between hierarchy and similarity. A closer look at our global patent map shows that the structure of the map reflects technological relationships across the hierarchical administrative boundaries of the subject matter specifications in the IPC scheme. While counts of IPC sections (i.e. the first letter of IPC codes, A, B, C, D, E, F, G, H) are commonly used as a measure of technological distance in patents, the 35 technological areas that are derived from cross-citations in our patent map often span multiple sections. For instance, the vehicles area includes six different sections, and the heating and cooling, construction and metals areas include five different sections. Textiles, lighting, semiconductors and chem and polymers include four different sections. Only medical devices, food, recording, computing and radio communication areas encompass a single section. This is strong evidence that the IPC on its own is not an appropriate framework to investigate technological diversity without taking technological distance into account. It is also a factor to consider in the analysis and definition of emerging technological fields and markets.

The degree of interconnectedness among technological categories adds another level of support to decision-making. Patent documents that reference other patents in similar technological areas have been suggested to offer incremental opportunities to advance an area, whereas patent documents that refer across diverse categories may offer the potential for radical innovation [34]. For instance, an interesting feature of our global patent map is the high level of interconnectedness of most of the 35 technological areas. This can be observed not only in many connections among technology groups within each technological area, as shown by the densest areas of the map, but also across them. Some exceptions are areas such as food, drugs and med chem, biologics, TV imaging and comm. cosm and med chem as well as radio and comm that form more uniform clusters of technology groups (i.e. they appear as clusters of nodes of the same colour) (Fig. 10.1). Another notable feature is the short distance among technologies in a handful of groups such as drugs and med chem and biologics, as shown by denser areas and darker lines in the left-hand side of the maps. The sparse areas of the map are those associated with technological areas that comprise fewer technological categories, including electric power, lighting and recording.

Finally, the concentration of patenting activity in the innovation landscape—or global map—is another aspect of visual analysis that can support strategic decision-making. Overlays created with patent data subsets allow this kind of analysis. Broader technology groups and more specific categories can be compared across organisations, and over time, to distinguish areas of R&D concentration and to identify trends, respectively. Areas of increasing activity can represent areas of market opportunity or be a signal of competitive threats when the analysis refers to specific companies and the purpose is to detect new entrants. Areas where technology development concentrates might anticipate emerging technological areas or niches. The complement, empty areas or 'white spaces' represent undeveloped areas. Our patent map, for example, uses categorisations to disaggregate some of the patent groupings into more fine-grained analysable components than other approaches. This more disaggregated clustering enables differentiation of the patent portfolios of, say, a company engaged in cosmetics patenting from one engaged in drug development and from yet another engaged in medical instrument development. Not shown here, but the maps can be blown up to allow closer examination of more fine-grained patterns.

Awareness of the conceptual heterogeneity of nodes or elements in the map raises the issue of whether the maps show 'similarity' among categories as we have assumed or other properties such as co-occurrence and complementarity. For example, patents of metals and automobiles are related not because these categories are similar but because automobiles are often made of metals. Also, plastics and metals may co-occur simply because they are materials that are used in similar products such as buckets and automobiles, not because they are similar. Moreover, unlike maps of science, where there has been a pre-established conventional understanding of disciplines, it is not straightforward how groups of technologies can be interpreted. This problem is compounded by the heterogeneous nature of the patent classes. Classes such as 'turbines and engines' include 'turbines' (F01D), 'jet propulsion' (F02K), 'aircraft equipment' (B64D) and 'airplanes and helicopters' (B64D). Elements from distinct branches of the IPC co-occur in maps, but rather than being similar, they likely co-occur because they are embedded and/or complementary. This difficulty that patent maps face is not simply a problem of classification, but a conundrum due to the multiple meanings and scales that the technology concept may take [35]. These issues suggest that the interpretation of patent maps should be ontologically flexible and one should take into account that both the elements and the relations may have different meanings.

A visual analytic study based on our base map and overlays would involve (1) creating a patent data set, (2) processing the data set to obtain overlay data by IPC-based category, (3) creating overlays and (4) analysing data and overlays to support decision-making. Users can draw on diverse data sources to create patent data sets for their analysis (e.g. we use EPO's PATSTAT). To process data sets, we have developed a mapping kit which includes source files that represent the structure of the base maps and thesaurus files that represent scientific publication and IPC-based category definitions and enable creation of overlay maps using software such as VantagePoint and Pajek.[3] The analysis typically involves comparing areas of concentration over time and across different entities such as companies, countries or technological fields. Overlays offer a general perspective that can be enriched with data tables with more detailed information on patenting activity.

10.6 Concluding Remarks

This chapter discusses broader aspects and issues of patent mapping in the context of information retrieval, exploration and analysis for patent corpora and their differences to similar approaches applied to scientific literature. The chapter also discusses visual analytics and diverse methods for mapping—including a new global patent map and overlay technique developed by the authors—that enable ‘visual thinking’ [13] and a better understanding of technology development concentrations and R&D profiles of companies or countries. To exemplify the kind of analytical support offered by global and overlay maps, we illustrate the application of the patent overlay maps we developed to benchmark the nanotechnology-related patenting activities of companies and reveal the areas of concentration of their patenting activities.

Patent analyses play key roles in competitive technical intelligence as they combine multiple information types to offer innovation indicators to support decision-making. Patent mapping provides a visual component to enrich various innovation system analyses and complements science overlay mapping to enrich our understanding of R&D activities. Patent overlay maps serve to discern patent activity patterns with implications for innovation. These can inform corporate investment decisions via intelligence about key competitors’ perceived trajectories, for example. Alternatively, analyses can contribute to policy discourse by profiling national positions and potential. Patent analysis involves data search and retrieval and a number of processes for data clean-up and refinement to obtain subsets that ultimately contribute insight into specific technologies or market sectors being engaged.

The visualisation of knowledge or technological landscapes has been a prominent part of publication and patent analyses since their origins, but recent improvements in computational power and algorithms have allowed the creation of diverse global maps of science and technology. Both science and patent maps draw on network analysis algorithms and visualisation procedures that help to explore and visually identify areas of activity, interrelationships and the overall structure of scientific and technological activities. Science and technology maps complement other methodological approaches to data analysis and can help to interpret and find

[3] This mapping kit is available upon request to the authors. VantagePoint is a commercial software for text mining: https://www.thevantagepoint.com. The purpose of the kit, however, is to make this mapping technique available for use with other software as well.

meaning in complex data by transforming abstract and intangible data sets into something visible and concrete. In patent maps, in particular, a key characteristic in being able to visualise innovative opportunities is the ability to graphically represent ‘technological distance’ or the extent to which a set of patents reflects different types of technologies. Scholars have pursued diverse approaches to scientific publication and patent record-level analysis to create global maps of science and technology. Consensus on the shape and relative position of science and technological categories are important to make global maps meaningful to compare, for example, organisational or technological patent data subsets.

Our patent mapping approach offers distinctive visualisation capabilities. In contrast to prior IPC-based global patent maps, our approach recombines IPC categories to reflect a finer distribution of patents. Thus, it enables improved differentiation ability in categories with a large amount of patenting activity. It also facilitates replication by helping to trace back individual categories to verify results and make improvements. One of the most interesting findings of our work is that IPC categories that are close to one another in the patent map are not necessarily in the same hierarchical IPC branch, which suggests that technological distance is not always well proxied by relying on the IPC administrative structure. The introduction of the Cooperative Patent Classification (CPC) scheme is likely to affect our category definitions or the process by which we come up with specific definitions. Still the overall dimensions of the map would be supported and only some of the topical areas in the margins would change.

Visualisations are valuable tools for competitive R&D and policy decision-making support. Potential applications of patent overlay maps include organisational and regional/country benchmarking (e.g. for the examination of competitive positions), exploration of potential collaborations and general analysis of technological changes over time. Patent maps may also reveal relatively unexplored technological areas that are more central to other technologies or highlight denser areas with more technological interdependency that might form platforms for the emergence of future technology applications. Ongoing work we undertake seeks to overcome some issues found in the development of the original patent overlay maps. The coverage of the technology classification scheme we developed is among the most important issues we address. While the data source may cover a wide range of IPC categories, new technologies and categories resulting naturally from innovation processes require constant updates to maintain good coverage and be able to support decision-making in emerging areas as well.

Acknowledgements This chapter is based on research undertaken at Georgia Tech drawing on support from the US National Science Foundation (NSF) through the Center for Nanotechnology in Society (Arizona State University, Award No. 0531194 and NSF Award No. 1064146) (‘Revealing Innovation Pathways: Hybrid Science Maps for Technology Assessment and Foresight’) in collaboration with the Center for Nanotechnology in Society, University of California Santa Barbara (NSF Awards No. 0938099 and No. 0531184). The findings and observations contained in this chapter are those of the authors and do not necessarily reflect the views of NSF.

References

1. Kay L, Newman N, Youtie J, Porter AL, Rafols I (2014) Patent overlay mapping: visualizing technological distance. J Assoc Inf Sci Technol 65(12):2432–2443. doi:10.1002/asi.23146
2. Rafols I, Porter AL, Leydesdorff L (2010) Science overlay maps: a newtool for research policy and library management. J Am Soc Inf Sci Technol 61(9):1871–1887
3. Porter AL, Newman NC (2005) Patent profiling for competitive advantage: deducing who is doing what, where, and when. In: Moed JF, Glanzel W, Schmoch U (eds) Handbook of quantitative science and technology research. Kluwer, Dordrecht, The Netherlands
4. Trippe A (2003) Patinformatics: tasks to tools. World Patent Inf 25:211–221
5. Dunphy SM, Herbig PR, Howes ME (1996) The innovation funnel. Technol Forecast Social Change 53:279–292
6. Souder WE (1987) Managing new product innovations. Lexington Books, New York, NY
7. Watts RJ, Porter AL (1997) Innovation forecasting. Technol Forecast Social Change 56:25–47
8. Balconi M, Breschi S, Lissoni F (2004) Networks of inventors and the role of academia: an exploration of Italian patent data. Res Policy 33(1):127–145
9. Schoen A, Villard L, Laurens P, Cointet J-P, Heimeriks G, Alkemade F (2012) The network structure of technological developments; Technological distance as a walk on the technology map. Presented at the STI Indicators Conference 2012, Montréal. http://sticonference.org/Proceedings/vol2/Schoen_Network_733.pdf. Accessed 22 June 2013
10. Ernst H (2003) Patent information for strategic technology management. World Patent Inf 25(3):233–242
11. Leydesdorff L, Kushnir D, Rafols I (2012) Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC). *Scientometrics* 1–17. doi:10.1007/s11192-012-0923-2.
12. Leydesdorff L, Alkemade F, Heimeriks G, Hoekstra R (2015) Patents as instruments for exploring innovation dynamics: geographic and technological perspectives on "photovoltaic cells". Scientometrics 102:629–651
13. Chen C (2003) Mapping scientific frontiers: the quest for knowledge visualization. Springer, London
14. Boyack KW (2003) An indicator-based characterization of the Proceedings of the National Academy of Science. Paper presented at the NAS Sackler Colloquium on Mapping Knowledge Domains
15. Archibugi D, Pianta M (1996) Measuring technological change through patents and innovation surveys. Technovation 16(9):451–468
16. Huang Y, Ma J, Porter AL, Kwon S, Zhu D (2015) Analyzing collaboration networks and developmental patterns of nano-enabled drug delivery (NEDD) for brain cancer. J Nanotechnol 6:1666–1676
17. Hinze S, Reiss T, Schmoch U (1997) Statistical analysis on the distance between fields of technology. Paper presented at the Innovation Systems and European Integration (ISE), Targeted Socio-Economic Research Program, 4th Framework Program of the European Commission (DGXII), Karlsruhe, Germany
18. Small H (1973) Co-citation in the scientific literature: a new measure of the relationship between two documents. J Am Soc Inf Sci Technol 24(4):265–269
19. Klavans R, Boyack KW (2009) Toward a consensus map of science. J Am Soc Inf Sci Technol 60(3):455–476
20. Bollen J, Van de Sompel H, Hagberg A, Bettencourt L, Chute R, Rodriguez MA et al (2009) Clickstream data yields high-resolution maps of science. PLoS One 4(3), e4803. doi:10.1371/journal.pone.0004803
21. Boyack KW, Börner K, Klavans R (2009) Mapping the structure and evolution of chemistry research. Scientometrics 79(1):45–60
22. Boyack KW, Klavans R, Börner K (2005) Mapping the backbone of science. Scientometrics 64:351–374

23. Janssens F, Zhang L, Moor BD, Glänzel W (2009) Hybrid clustering for validation and improvement of subject-classification schemes. Inf Process Manag 45(6):683–702. doi:10.1016/j.ipm.2009.06.003
24. Leydesdorff L, Rafols I (2009) A global map of science based on the ISI subject categories. J Am Soc Inf Sci Technol 60(2):348–362
25. Moya-Anegón F, Vargas-Quesada B, Chinchilla-Rodríguez Z, Corera-Álvarez E, Herrero-Solana V (2007) Visualizing the marrow of science. J Am Soc Inf Sci Technol 58(14):2167–2179
26. Moya-Anegon F, Vargas-Quesada B, Herrero-Solana V, Chinchilla-Rodriguez Z, Corera-Alvarez E, Munoz-Fernandez FJ (2004) A new technique for building maps of large scientific domains based on the cocitation of classes and categories. Scientometrics 61(1):129–145
27. Rosvall M, Bergstrom CT (2010) Mapping change in large networks. PLoS One 5(1), e8694. doi:10.1371/journal.pone.0008694
28. Van den Besselaar P, Leydesdorff L (1996) Mapping change in scientific specialties: a scientometric reconstruction of the development of artificial intelligence. J Am Soc Inf Sci Technol 46(6):415–436
29. Breschi S, Lissoni F, Malerba F (2003) Knowledge-relatedness in firm technological diversification. Res Policy 32:69–87
30. Jaffe A (1986) Technological opportunities and spillovers of R&D: evidence from firms' patents, profits, and market value. Am Econ Rev 76(5):984–1001
31. Kauffman S, Lobo J, Macready WG (2000) Optimal search on a technology landscape. J Econ Behav Organ 43:141–166
32. Rafols I, Leydesdorff L (2009) Content-based and algorithmic classifications of journals: perspectives on the dynamics of scientific communication and indexer effects. J Am Soc Inf Sci Technol 60(9):1823–1835
33. Boyack KW, Klavans R (2008) Measuring science–technology interaction using rare inventor–author names. J Informet 2:173–182
34. Olsson O (2004) Technological opportunity and growth. J Econ Growth 10(1):35–57
35. Arthur WB (2010) The nature of technology. Penguin, London