World Patent Information 55 (2018) 1–18

Contents lists available at ScienceDirect

World Patent Information

journal homepage: www.elsevier.com/locate/worpatin

Tech mining to validate and refine a technology roadmap

Geet Lahoti[a,b,*], Alan L. Porter[c,d], Chuck Zhang[a,b], Jan Youtie[e], Ben Wang[a,b,f]

[a] School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[b] Georgia Tech Manufacturing Institute, Georgia Institute of Technology, Atlanta, GA, USA
[c] Technology Policy & Assessment Center, Georgia Institute of Technology, Atlanta, GA, USA
[d] Search Technology, Inc. Norcross, GA, USA
[e] Enterprise Innovation Institute, Georgia Institute of Technology, Atlanta, GA, USA
[f] School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA

ARTICLE INFO

Keywords:
Technology roadmap
Tech mining
Scientometrics
Patent analysis
Nanocomposite coatings

ABSTRACT

This study uses 'tech mining' (extracting intelligence from R&D data) to validate and refine the content of a particular section of a landmark roadmap of nanotechnology for aeronautics. We utilize topical content from publications and patents to analyze the developmental status of nanocomposite coating technologies. This enables us to validate predictions made by specialists, as presented in the target technology roadmap section. Moreover, we augment that roadmap section by providing additional information on nanocomposite-related emerging technologies. This study supports use of tech mining as a means to inform technology roadmapping, both when creating a new roadmap and to check progress.

1. Introduction

Technology roadmaps are considered one of the most influential tools for technology management [1]. Roadmapping assists in developing and implementing strategic business and product plans as it provides the information, processes and tools needed to produce comprehensive plans [2]. It acts as a tool for depicting the evolution of technology, which helps in predicting its future [3]. It can serve specific objectives of firms and governments. Graphical representations provided by technology roadmaps help firms gain understanding of the relationships among markets, products and technologies over time [4,5]. Technology roadmaps help governments and industry make investment decisions efficiently and remain competitive by providing intelligence on emerging technologies [1].

Depending on the intended purposes and graphical formats, there can be several types of technology roadmaps [2]. Among the several classifications as proposed by researchers, the most popular one divides a large number of roadmaps into four groups: science-technology roadmaps, product-technology roadmaps, industry roadmaps and product roadmaps [6]. Although such roadmaps can respond to the needs of different organizations, most practitioners find it difficult to develop a technology roadmap using traditional methods based on individual experiences and domain knowledge [7,8]. Qualitative methods, such as Delphi-based technology roadmapping, involve integrating diverse opinions of experts retrieved through surveys and questionnaires. These methods are time-consuming and costly and suffer from low response rates [9]. Since these methods depend on experts' intuitive knowledge and opinion, which may be influenced by subjective elements and limited cognitive horizons, they may be biased [10–12]. Also, we hypothesize that qualitative information used in developing technology roadmaps are rarely validated with quantitative information. Tech mining techniques could provide such validation, but are not widely used this way. Researchers have been looking for quantitative technological information in patents, academic papers, technology standards, and technical reports using systematic techniques involving data mining [1].

Data mining can uncover interesting patterns in databases that assist in decision-making [13]. It allows researchers to retrieve information from a variety of data formats or databases. Text mining extracts information from text-based databases, which contain collections of unstructured or semi-structured documents, often in the form of abstracts of research papers, reports, patents and so on [1]. It is also a tool to identify themes running through documents and summarizing their contents [14], explore technological opportunities [15], cluster similar documents [16] and classify newly generated documents [17]. Analysts often use keywords to represent the content of a document and exploit them for knowledge discovery [18].

'Tech mining' is the process of deriving technology intelligence from science, technology and innovation (ST&I) information record sets through text mining [19]. The main purpose for this technique is to

* Corresponding author. School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA.
E-mail address: glahoti6@gatech.edu (G. Lahoti).

https://doi.org/10.1016/j.wpi.2018.07.003
Received 4 February 2018; Received in revised form 28 March 2018; Accepted 7 July 2018

explore research and development (R&D) activities in publication and patent record sets. Over the past 25 years, Technology Policy and Assessment Center at Georgia Institute of Technology has been developing this technique and its variants to analyze ST&I resources for understanding changing and emerging technologies [15,20]. These techniques aid in realizing newly emerging science & technology (NEST) competitive intelligence (CI) [22,23,80]. Special attention has been paid to the development of 'term clumping' steps which are required for cleaning, consolidation and clustering of terms and phrases obtained by applying Natural Language Processing (NLP) to the ST&I records [24]. These terms and phrases are indicative of significant topics related to the technologies treated in the research papers and patents. While dealing with a huge set of ST&I records, one may face thousands of similar terms and phrases. In scenarios such as this, 'term clumping' steps can assist in obtaining topical intelligence. This can help expedite the process of technology roadmapping [3,23,25] and advance efforts to use the Forecasting Innovation Pathways [26–28] method.

This paper uses tech mining in a validation mode to analyze a particular section of the well-regarded 2012 NASA Nanotechnology Roadmap. In 2010, the National Aeronautics and Space Administration (NASA) identified fourteen technology areas and asked the agency specialists, who were assigned to each of these areas, to define the respective technology area's top technical challenges and their impact on spaceflight missions [http://www.nasa.gov/offices/oct/home/roadmaps/index.html]. In December 2010, a set of developed draft roadmaps covering the near-term space technology needs, longer term challenges, NASA's space technology activities prioritization, and overall technology investment strategy, was made publicly available and presented to the National Research Council (NRC) for critique. The final version of the roadmaps published in 2012 contains the original draft roadmap along with NRC's comments. NRC also provided NASA with recommendations on future technology investments. NASA mentioned in the roadmaps that they plan to update them on a regular basis because forecasting the wide range of possible advancements in the various technology areas is very difficult.

Considering the fact that NASA involved agency specialists, who can be treated as experts, in developing the roadmaps and contracted with NRC to get recommendations, we can say the roadmapping process seems to be highly based on experts' opinions. As outlined earlier, this qualitative way of developing roadmaps may suffer from several disadvantages. Moreover, NASA stated that they intend to revise those roadmaps on a regular basis. This motivates us to undertake this study where we consider a roadmap – Engineered Materials and Structures Roadmap (excerpted from Nanotechnology Roadmap, Technology Area 10, National Aeronautics and Space Administration, April 2012) as shown in Fig. 1 and use tech mining to validate its content and refine it. The chosen roadmap has several themes: lightweight structures, damage tolerant systems, adhesives, coatings, and thermal management. In this study, we just consider the theme "coatings." As mentioned in the roadmap (Fig. 1), this technology is capable of contributing to "extreme environment operations," "efficient extravehicular activity (EVA) operations," "thermal control/protection," and "smart airframe and propulsion." Also, this technology is expected to mature by 2017-26 according to this roadmap. More specifically, durable tailored nanoengineered surfaces and nanostructured coatings are expected to mature by 2017-26 and 2020-26 respectively. We intend to validate the emergence of these technologies by exploring R&D and innovation activities in those areas using tech mining. In particular, we mine publication and patent records to derive topical intelligence. We analyze the trends of the related topics and try to find out what supports the experts' predictions. Further, we aim to refine this roadmap by providing additional information that we obtain during the publications and patents topical analyses. On the one hand, publications are a good source of technological information on state-of-the-art research, theoretical development, and potential technologies related to a particular research area [1]. On the other hand, patents provide information on promising technologies and their practical characteristics. Therefore, the trend analysis of publication and patenting activities could help us ascertain bases for the predictions presented in the technology roadmap section. During this process, we may come across some additional information that will help in refinement.

There are six parts in this paper. This part gives a general introduction covering technology roadmaps, tech mining, and the technology roadmap under consideration – the NASA Nanotechnology Roadmap. In the next part, related work is reviewed. The third part details the dataset we used and our methodology for deriving topical content related to nanocomposite coatings. In the fourth part, we present the results of our nanocomposite coatings topical analysis. Developmental trends of the nanocomposite coatings related technologies obtained from topical analysis are also discussed in this part. Validation and refinement of the technology roadmap is discussed in part five. Last, we offer a conclusion, along with challenges, limitations, and a scope for future work.

2. Literature review

Researchers have been advancing the development of technology roadmapping procedures in three directions: qualitative methods, quantitative methods and a combination of both. There is no 'official guidebook' for developing technology roadmaps [3]. Qualitative methods such as Delphi, scenario planning, expert interviews, and discussion/seminars/workshops are among the most prominent methods for roadmap development and implementation [2,29]. Beginning with informal discussions, one roadmapping process becomes a series of dedicated workshops, involving industry stakeholders, government officials and/or researchers [30]. The 3-phase process for roadmapping – preliminary activity, roadmap development, and follow-up activity, as developed by Sandia National Laboratories [31]; [29], and its variants were used for a variety of emerging technologies such as microsystem and nano-system [32], semiconductor silicon industry [33], and pharmaceutical technology [34] [35]. developed 'T-Plan' which supports the swift initiation of roadmapping in three stages: planning, roadmapping and roll-out [36]. introduced a modified T-Plan process with five key modules. Researchers have contributed to construct general schemes so that wider strategic needs could be fulfilled. Modularization method-based customizable technology roadmapping function [37] and catalog for technology management-oriented analytics [38] were some of the efforts in that direction. Aiming to construct a flexible toolkit for supporting strategic technology management [39], proposed a workshop-based approach to address a range of management challenges.

Responding to the needs of data-driven roadmapping, there has been substantial research into the quantitative approach to technology roadmapping. Since the late 1990s, text mining and bibliometrics have been progressively used to extract useful information from ST&I records [40]. Among several ST&I text analysis techniques, retrieving topics via textual elements (e.g. words, terms, or phrases) and then identifying relationships via defined association rules is widely popular [41]. Although several quantitative methods for information extraction and visualization have been developed, they could not be explicitly termed as quantitative roadmapping methods but they implicitly support roadmapping processes. A semi-automatic approach for knowledge discovery in bibliographic databases [20], a generalized approach for detection of emerging trends from co-citation networks and their visualization [42], and a unified approach to mapping and clustering of the most frequently cited publications [43] were some of the studies in that direction. Also, researchers have been exploiting statistical techniques for finding new events in a stream of broadcast news stories [44] and retrieving linguistic relationships from patents and discovering technological trends [45]. In an attempt to develop a more effective roadmapping procedure [46], proposed a methodology, which made use of information extracted from product manuals and patent