21ST INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED17
21-25 AUGUST 2017, THE UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, CANADA

A VISUAL ANALYSIS OF TECHNICAL KNOWLEDGE EVOLUTION BASED ON PATENT DATA

Smojver, Vladimir (1); Potočki, Eva (2); Štorga, Mario (2)
1: Vehicle Center of Croatia, Croatia; 2: University of Zagreb, Croatia

Abstract

Knowledge visualization can provide a narrative for understanding the dynamics of knowledge dissemination, synthesis and redistribution. The interactions between various knowledge instances is not cumulative, but spawns new knowledge instances. These knowledge instances can then spawn new knowledge instances. This is a process which cannot be controlled, but its dynamics can be understood through visualization. Using visualization, patterns of knowledge asset interactions that can lead to innovation and success, or conflict and failure, can be identified. In today's knowledge economy, organizations that operate in highly competitive market conditions have a need for up-to-date knowledge of emerging technologies. This research attempts to provide insight into a technology fields evolution by analysing how the content of patent claims changes over time. This is achieved by analysing keywords that occur in patent claims and using this data to create topics of interest addressed by patents.

Keywords: Knowledge management, Patent map, Innovation, Technology evolution, Technology

Contact:
Vladimir Smojver
Vehicle Center of Croatia
Vehicle Testing Department
Croatia
vsmojver@gmail.com

Please cite this paper as:
Surnames, Initials: *Title of paper*. In: Proceedings of the 21st International Conference on Engineering Design (ICED17), Vol. 6: Design Information and Knowledge, Vancouver, Canada, 21.-25.08.2017.

1 INTRODUCTION

Knowledge visualisation can provide a narrative for understanding the dynamics of knowledge dissemination, synthesis and redistribution. The interactions between various knowledge instances are not cumulative, but spawn new knowledge instances. This is a process which cannot be controlled, but its dynamics can be understood through visualisation. Using visualisation, patterns of knowledge asset interactions that can lead to innovation and success, or conflict and failure, can be identified. (Štorga et al., 2013) In today's knowledge economy, organisations that operate in highly competitive market conditions have a need for up-to-date knowledge of emerging technologies (Dedehayir et al., 2014).
In the research presented in this paper, an attempt is made to visualise the evolution of technical knowledge by means of the inventions recorded in patents. Patents are used as proxies for technical innovation as explained in previous work (Smojver et al., 2016).
Examining a technologies evolution can provide invaluable insight into the implicit rules that technologies follow throughout their life cycle. A close examination of a knowledge evolution with certain technology field can provide information about technology current life cycle status as well as its distribution and impact. Finally, exploring a technologies evolution might be used in simulations of possible future directions of the technology development which could reduce the uncertainty of decision-making in new product development projects. The visualisation of invention evolution captured in patent data provides the usual benefits of visualisation techniques; it provides a means to understand data more intuitively and, in some cases, enables the implementation of quantitative analysis techniques (e.g. if complex network analysis is applied).
Despite the growing number of studies related to the visualisation and analysis of the patents within technology field, there is very limited research that uses patent content and context information to conduct a dynamic analysis of the technology evolution over time. Therefore, this research attempts to provide insight into a technology fields evolution by analysing how the content of patent claims changes over time. This is achieved by analysing keywords that occur in patent claims and using patents' metadata to discover the topics of interest addressed by patents.
This rest of the paper is structured as follows: first, the theoretical background of using patent data in the field of technology management is given. Then, an overview of existing patents maps provided by the JPO (Japanese Patent Office) and WIPO (World Intellectual Property Organization) is provided. Next, a literature review covering the most popular types of patent maps developed by researchers is conducted. Finally, a case demonstrating the use of a dynamic patent map is presented.

2 LITERATURE REVIEW

2.1 Patents

A patent is a formalised document that contains information about technical inventions and is an invaluable asset that can be utilised in a myriad of areas such as R&D, technology management, financial investing, law and economic analysis. Patent information is a useful indicator for the technological development strategies or global strategies of individual enterprises in response to intensifying competition (Suzuki, 2011). If carefully analysed, they can show technological details and relations, reveal business trends, inspire novel industrial solutions or help make investment policy (Campbell, 1983). Since their structure is uniformed, they have been proved to be a resource open to automatic data retrieval and analysis, such as data mining and text mining (Yoon ad Park, 2004a; Tseng et al., 2007).
A drawback when analysing any large data set is the inability to gain intuitive insight based on the data alone. Visualising data is a way to solve this problem by providing a clear representation of examined data which is more easily understood by non-experts such as management or other miscellaneous stakeholders.
Since the analysis of patent data usually involves examining a very large data set, the visualisation of this data can augment understating of the information contained in patents and its context. Visualised expression of patent analysis results, which enables understanding of a various patent information complexity easily and in an effective manner, is called a patent map (http://www.wipo.int). The patent map is usually created by gathering patent information within a target technology field, processing, and analysing it. Patent maps are generally defined as "a visual form – a chart, table or graph – that analyses or arranges patent data to make to more informative and constructive" (Lee et al., 2009).

2.2 Patent Maps

Suzuki defines patent maps as “patent information collected for a specific purpose of use, and assembled, analysed and depicted in a visual form of presentation such as a chart, graph or table” (Suzuki, 2011). The World Intellectual Property Organization and particularly Japan Patent Office provide a few standard types of patent maps that stakeholders can integrate into their development and research. These standard types of patent maps are listed in Table 1.

Table 1. Representative Examples of Patent Map Types provided by WIPO and JPO

WIPO	*JPO*
Rate Map	Element-Based map
Number Map	Diagram of Technological Development
Trend Map	Interpatent Relation Map
Relation Map	Matrix Map
Radar Map	Systematized Art Diagram
Portion Map, etc.	Time Series Map
List Map	Twin Peaks Analysis Map
Matrix Map	Maturation Map
TEMSET Map	Ranking Map
Development Map	Share Map
Problems Vs. Solutions map	Skeleton map

The definition of patent maps provided by IP organisations is expanded by researchers who differentiate between the visualisation of the structured items (for example patent number, filing date or investors) and the unstructured items (patent text, abstract, claims). The visual representation of the former is usually called *patent graphs* by researchers, and those of the latter are called *patent maps*. It should be noted the term “patent map” can loosely be used in both of these cases (Kim et al., 2008). Even though IP organisations provide representative examples of different kinds of patent maps, research in this field is being conducted with the aim to expand and enhance current types as well as introduce new types of patent maps. Common approaches to patent maps found in literature can be summarised as presented in Table 2 and are described in more details in the following subsections.

Table 2. An overview of methods from literature

Author	**Input data**	**Analysis Method**	**Output**
Yoon et al. 2004	Patent citations	Network analysis	Citation map and quantitative analysis
Suh et al. 2009	Extracted keyword	k-Means clustering	Semantic network of keywords sorted by filling date and frequency in patent documents
Lee et al. 2009	Extracted keywords	PCA used to reduce number of keyword vectors suitable for 2D map	Patent „vacancies are identified
OuYang 2011	Patent citations, Extracted keywords	Patent family, TRIZ	Key patents, Technology performance map
Yang et al. 2012	Extracted keywords	Dependency tree generated by parser	Conceptual graph
Son et. al. 2012	Extracted keywords	Generative topographic mapping	Automatic detection and interpretation of patent vacuums
Su et. Al 2009	Patent citations	Network analysis	Citation map and quantitative analysis
Yoon et al. 2002	Extracted keywords	Self-organizing feature map	Technology vacuum map Claim point map Portfolio map
Tseng et al. 2007	Extracted keyword	Multi-stage clustering	Topic Map

2.2.1 Citation-based patent maps

The most common topic in the research are patent maps that visualise a citation based network between patents. Such network is represented as a directed graph where nodes represent individual patents and edges represent citations between them. Yoon and Park (2004b) propose a network-based analysis for patent citation analysis. OuYang et al. (OuYang and Weng, 2011) combine patent families with patent citation analysis in a new product design (NPD) process calling the approach the New Comprehensive Patent Analysis approach for NPD. They identify key patents for the purpose of product development design and use TRIZ theory to analyse the technological performance of patents. Su and Lee (2009) create a patent citation network and use network properties (degree centrality, betweenness centrality, closeness centrality) to calculate technology evolution mechanisms (technology convergence, technology diffusion etc.). Smojver et al. (2016) propose a dynamic citation based patent map and propose the network growth rate over the time as a mean to determine the current life cycle stage of technology. Network based patent maps can also be categorised according to a specific purpose: the technical patent map, the management patent map, and the claim patent map (Yoon et al., 2002).

2.2.2 Keyword based patent maps

Another common approach for visualising patent content related information is creating a keyword based patent map. The keywords are primarily extracted from the title of the patent, abstract or claim using some form of text processing. Such patent maps usually show these keywords clustered together based around the topic or function. Kim et al. (2008) use the k-Means algorithm to cluster patent documents by analysing the collected keywords of a targeted technology field. Lee et al. (2009) propose an approach for creating and utilising keyword-based patent maps for use in new technology creation activity. Text processing is used to transform patent documents into structured data in order to identify keyword vectors. Principal component analysis is employed to reduce the number keyword vectors to make suitable for use in a two- dimensional map. Finally, patent "vacancies", defined as blank areas in the map that are sparse in patent density but large in size, are identified.

Yang and Soo (2012) developed a technique to extract conceptual graphs from a patent claim using syntactic and semantic information. After information extraction, the patent claims sentences can be mapped into a single connected conceptual graph. A conceptual graph is a kind of formal knowledge representation in terms of semantic networks and existential graphs. Yoon et al. (2002) propose and exploratory process of developing a self-organizing feature map (SOFM) based patent map that visualises the complex relationship among patent and the dynamic pattern of technological advancement. They suggest three types of patent maps: technology vacuum map, claim point map and technology portfolio map.

Suh and Park (2009) propose a three-dimensional visualisation method and analysis tool based on keywords which contributes to evaluating emerging technologies for services. It calculates and displays five values of a keyword: level, time, the relative growth the relative share, and the relative order. Son et al. (2012) propose a generative topographic mapping (GTM) – based patent map, which aims to automatically identify a patent vacuum. This approach overcomes any subjective influences in detecting patent vacuums since it creates a grid-based two-dimensional map in which each patent is mapped into the relevant grid. An empty grid, therefore, denotes a patents vacuum.

In this paper, a dynamic patent map based on visualising the dynamics of keyword creation in patent claims is presented.

3 RESEARCH METHOD

In here presented research, a dynamic (evolutionary) keyword based patent map is created. Keywords are used to track the evolution of technical knowledge captured within patents. The results of the keyword evolution analysis enable the identification of technologies utilised in patents at a certain time point and augment our understanding of technologies that contributed the most to the development of the technology area. By analysing formed groups of keywords – clusters – insight is given regarding the topic development within the technology area.

NLP (Natural Language Processing) techniques (http://www.alchemyapi.com/) were used to extract the keywords from individual patents within the technology field. Manual filtering and normalisation of the extracted keywords was performed on such extracted information. The most relevant keywords were

selected for every patent, based on the score provided by the NLP engine and a keyword based network was created applying the rule if two keywords (nodes) belong to the same patents, the link (edge) should be created between them. The OgranicViz (http://www.organicviz.org/) tool was used for visualisation and dynamic analysis of the created network. The result of this is the formation of a keyword cluster communities that have more influential keywords in their respected centres. The publication date of the patent enables a dynamic analysis of the knowledge evolution and identification of the topics change over the time. The analysis of the network growth dynamics was applied in order to understand the main trends and identify key most influential keywords describing the technical knowledge.
As a cases study in the presented research, safety ski binding patents, IPC (International Patent Classification)classification code A63C9/08, are analysed. Prior research on created a patent citation network and used network analysis how the number of citations can be used to determine a technology field life cycle stage. Moreover, other standard methods to analyse the patent from a certain technology field were applied. One of the results of this research was the identification of the 4 points in time were the technology area transitions from one life cycle stage to the other (Smojver et al., 2016)

4 CASE STUDY RESULTS

The evolution of the knowledge
The results of the dynamic analysis of the technical knowledge network for the case study were observed in 4 key moments in time which correspond to the 4 life cycle stages of the examined technology previously detected: emergence 1965-1968, growth 1969-1975., maturation 1976-1978., and saturation 1979-2015 (Smojver et al., 2016). These stages were determined by conducting a citation analysis of the retrieved patents as opposed to this paper which analyses keywords that occur in patent claims. At the end of the first phase, there were only three patents. There was no overlap between the keywords of the individual patents so there are no links between the three groups of keywords (Figure 1).

During the second phase of the technology life cycle, there was a significant growth of the keyword network reflecting the nature of this phase which is characterised by rapid adding of the new knowledge and booming in patent applications. Most of the key patents for this technology occur in this phase, therefore the keyword network at the end of the phase is a lot more complex and topic clusters can be discerned. Figure 2 shows the clusters visually separated and marked in different colours. Key patents can be identified in the centres of larger clusters which demonstrates their importance for the observed technology. These star-shaped network communities demonstrate the existence of a popularity bias, meaning nodes with more connections have a higher probability of new connections.

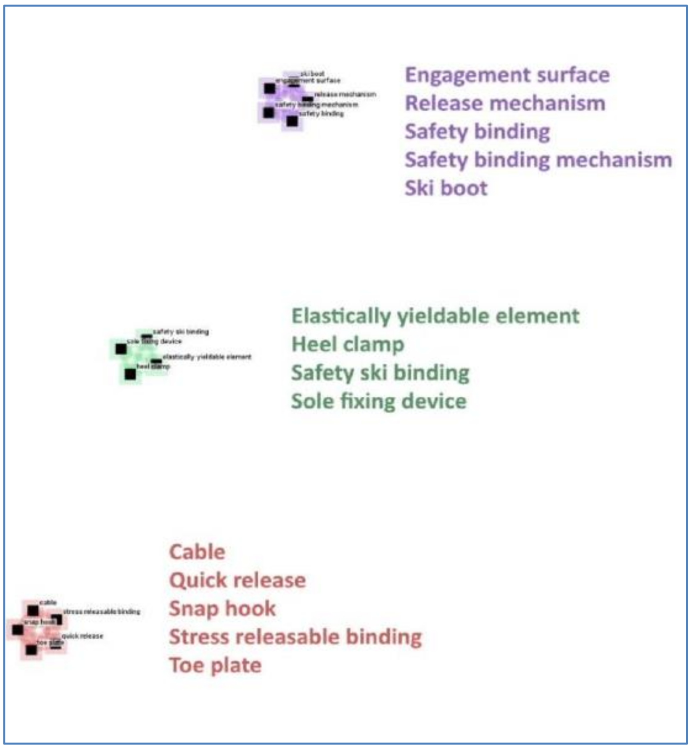

Figure 1. Emergence phase TLC

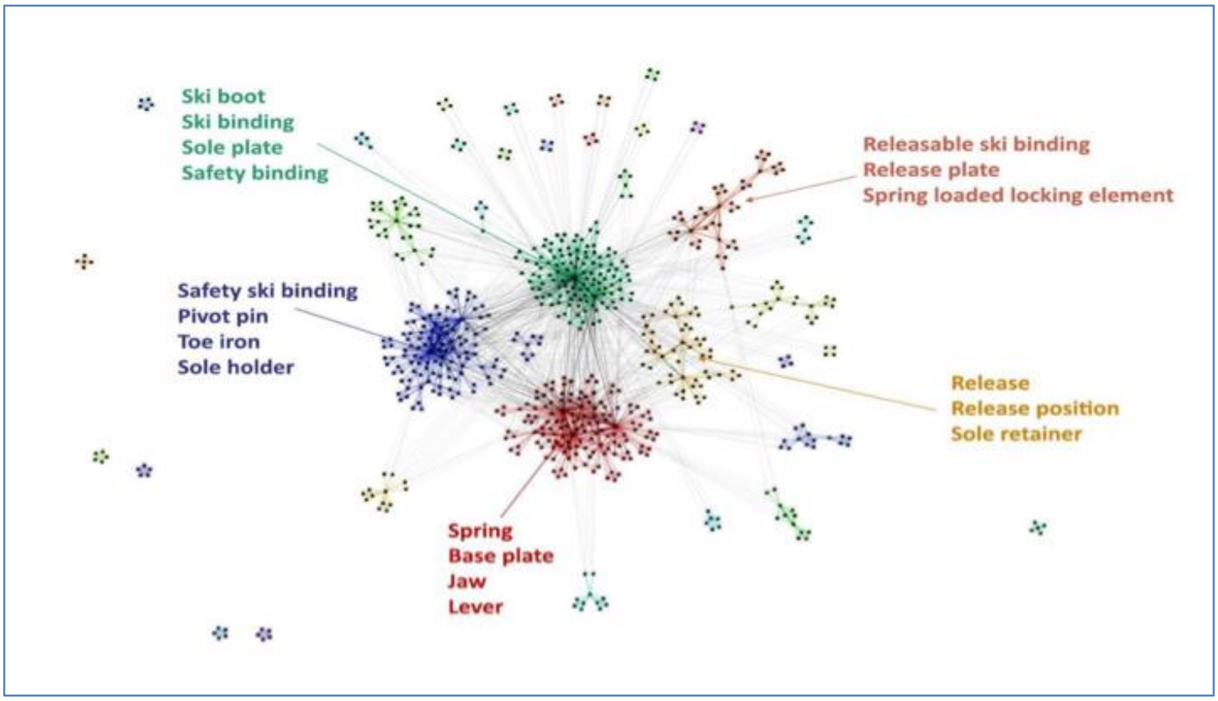

Figure 2. Growth phase TLC

During the final two phases, the maturation and saturation, the network became denser and restructured. The number of edges between keywords increases indicating the nature those phases where the proposal of the new solutions slowed down, and mainly refinement of the previously proposed solutions happened. In order to gain a better insight into the state of technology, individual clusters were examined.

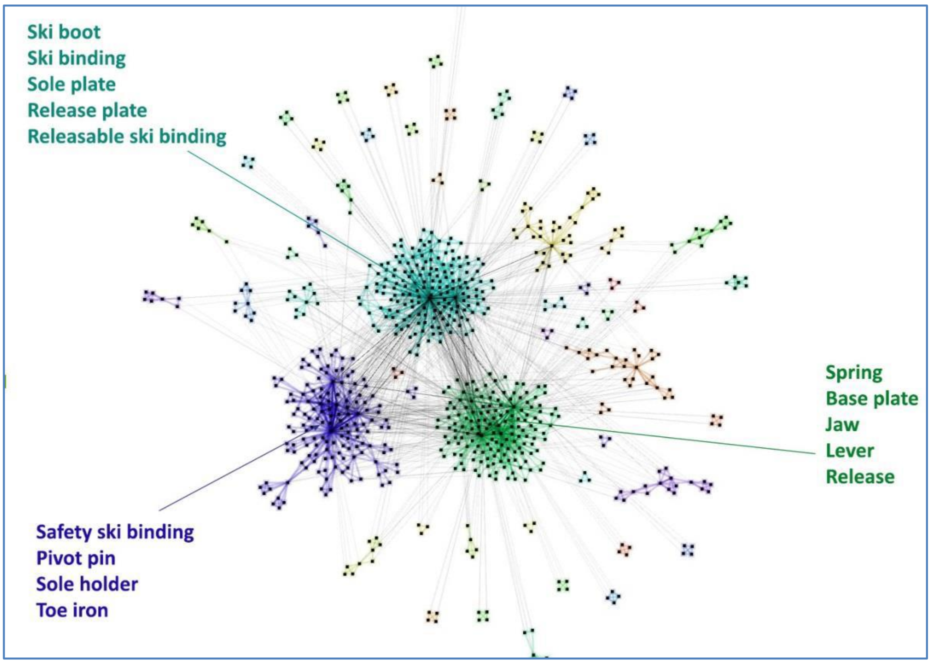

Figure 3. Maturation phase TLC

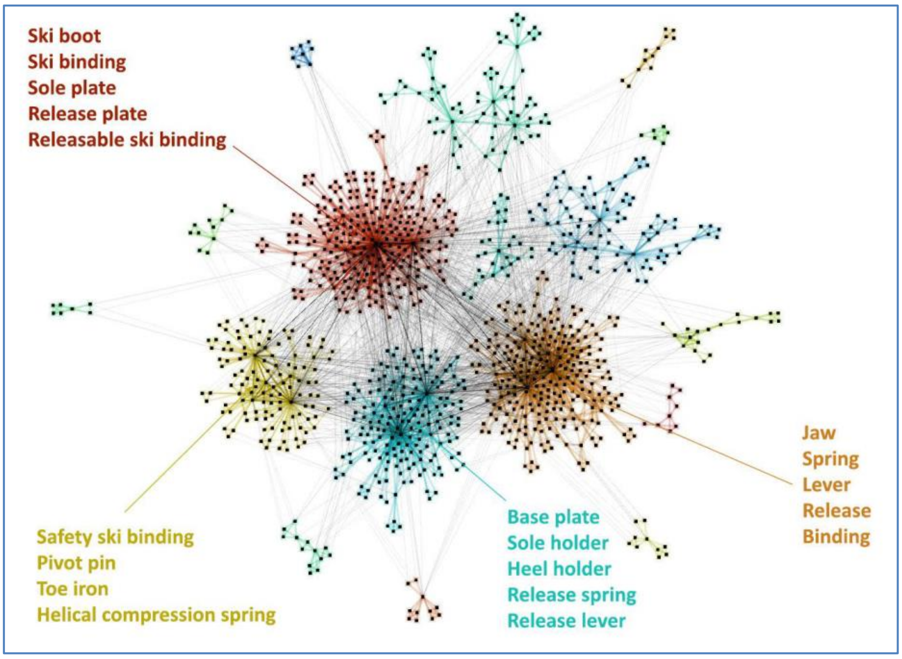

Figure 4. Saturation phase TLC

From this analysis of the keywords clusters in more details, it can be deduced that patents in this technology field separately focusing on two components of the ski binding: the front part (jaw) and the rear part (sole). Moreover, it is observed that most recent patents enable the release of the ski bind with a system of springs and elastic elements when the system is overstrained.

The evolution of topics

In the next step of the analysis, the keyword clusters were labelled as thematic topics and evolution of the topics were analysed (Figure 5).

Topic Number	Topic	Emergence 1960.–1968.	Growth 1969.–1975.	Maturity 1976.–1978.	Saturation 1979.–2015.
T1	Cable binding with safety release	5	9	9	11
T2	Safety ski binding with elastically yieldable element	5			
T3	Safety ski binding with a longitudinal release mechanism	5			
T4	safety locking mechanism achieved by levers and a mechanism with springs and elastic elements		138	150	249
T5	Adjustable parameters of the ski binding		43		
T6	Upward heel release mechanism		52	11	
T7	Heel hold down device		13		
T8	Lateral release of heel binding mechanism		21	26	
T9	Magnetic holding for ski binding		28		
T10	A system for fixing the boot toe during use		12	12	12
T11	Different principles of ski boot acceptance/release		120	167	261
T12	Using an electric signal for bind release		7		
T13	Elastic safety strap to prevent ski drift after release		11	11	10
T14	Binding is achieved with one connection consisting of complementary parts on the ski and boot		15	18	13
T15	Compression springs forces adjustments and pre stressing		114	143	120
T16	Different methods for automatic bind release			32	
T17	A mechanism for the simultaneous release of bind in two axes			6	
T18	A plate that supports and immobilises the heel of the boot that releases the ski in case of a fall				10
T19	A mechanism for locking the rear part of the bind when coupling with the boot				11
T20	An additional electrically controlled mechanism for ski binding release				19
T21	Release mechanism of toe piece during fall				76
T22	New mechanisms for ski boot holding				30
T23	New solutions including sensors and other electronic and digital elements				178
T24	Ski bindings releasable by ski poles				91
T25	Breaking mechanism for ski				8

Figure 5. The evolution of thematic topics for safety ski bindings

The groups were marked with *Tn* in order to observe the occurrence of new topics and recurrence of topics more easily (*n* denotes the ordinal number of the topic). Figure 5 shows the emergence of all of 25 identified topics over the time. The numbers on the timelines show the number of keywords the respected topic contains at a certain phase.

In order to further understand topics evolution dynamics, the content network growth analysis was conducted by the method proposed by Cash et al. (2013). Figure 6 shows the results of a network growth analysis where the keyword network is generated and continuously recalculated whenever a new keyword is added. This allows for an animated visualisation of the network over time and gives insight into the dynamics and evolution of the network as new keywords are added. The positive trends in growth analysis are corresponding with the dominant addition of the new nodes, in this case keywords, into the network, and negative trends with dominant rewiring of the existing nodes. The rate of growth δ if the network's edges in respect to network's nodes over time can be analysed by using methodology proposed by Cash et al. (2013).

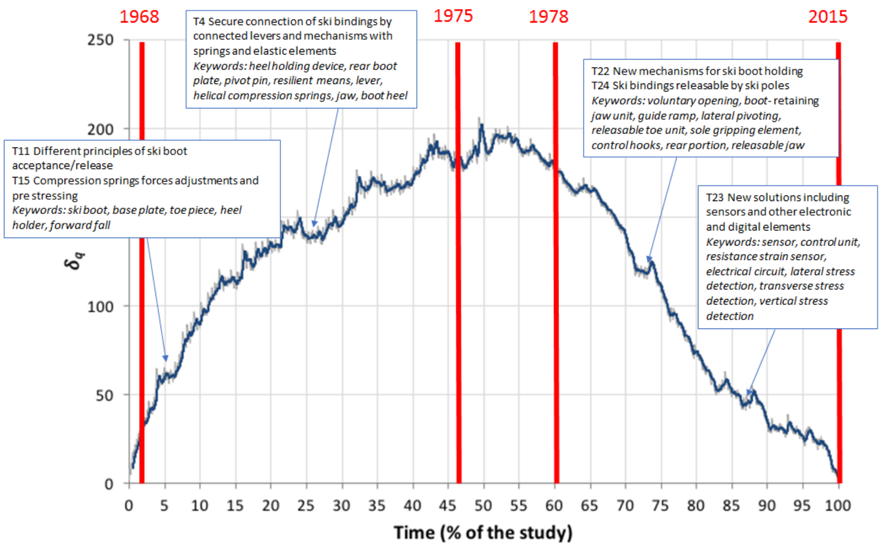

Figure 6. Content network relative growth analysis

In the first part of the study (0-50 % of the total time), more new keywords are introduced than patents are applied, meaning new patents, in average, introduce new keywords and positive network growth rate. The topics identified in this period represent dominant technical solutions that were most used in the creation of new patents. In the second part (45-100 % of the total time), there are more new patents then there are new keywords causing negative network growth rate. The meaning is that new patents applications mostly used existing keywords as a decline the innovativeness of new inventions. Several peaks should be noted in this period which denote the appearance of innovative patents which introduce new keywords and thus topics (at around 73% and 87 % of the total time). These peaks represent the introduction of new technologies that proved influential in the development of the field. These points in time represent crucial points in the technologies development where a significant change occurred. The topics identified at the end of the study (87% of the total time) might represent new technologies and could provide insight into the future development of the technology field.

5 DISSCUSION

In this paper, an approach is presented for creating a dynamic patent map based on visualising the dynamics of keyword creation in patent claims. Several key contributions could be further discussed.

First, the importance of data visualisation is given and the advantages of its application to patent analysis are listed. By visualising how patents interact in a dynamic setting, insight can be gained into the evolution of a technology field based on how patents of inventions in this technology field interact through time.

A method for dynamic keyword based patent map analysis is presented. The keywords are used to track the contextual evolution of patents based on a specific technology area. The results of the keyword analysis are used to identify technologies that were used in patent applications at certain points in time and which technologies contributed the most to the development of the technology field. Based on the formed communities, insight is gained onto the thematic development of the field and how the technology evolves over time.

Finally, this method of technology invention evolution trend analysis provides a base for understanding the trajectories of knowledge evolution. These visualisations are helpful, not only to engineers and data analysts, but also other stakeholders (managers, investors, customers) which only adds to their value.

To summarise, this research introduces a method to observe the evolution of keywords used in patent applications and uses this data to identify topics of interest at the certain point of the technologies evolution. These topics provide insight in the direction of research in the technology field as well as what developed technologies were most influential on future research. Based on these results, insight may be gained in the future development of an observed technology field. A limitation inherent to this method, as well as all patent based methods, is the uncertainty that all relevant inventions have been patented as well as that the data set comprising of retrieved patents is complete. However, the volume of retrieved patents is large enough that a general trend of development can be surmised.

6 CONCLUSIONS

Recognising the patterns that occur in technology invention evolution enables the creation of models that can be used to predict future development which may prove to be an invaluable resource to stakeholders. The analysis of technological development is increasingly important in the early stages of design where possible solutions to subsystem requirements are being considered. A systemised visualisation of technologies applied up to date may contribute to the creation and development of an optimal solution. Thus, by eliminating solutions that are proven to be ineffectual, the time to create concepts, and thus the entire design time, is reduced.

There are several issues that should be mentioned as well. An issue frequently brought up when conducting any type of patent based research is the validity of patents as a data source. A popular argument is that not all inventions are patented, therefore patents are not reliable when trying to evaluate a technological are in any capacity. While the fact that not all inventions are patented is true, the general body of patent applications is large enough that trends can be discerned. Moreover, the aim of this research is to see technology evolves within the scope of patents as proxies for technology invention. Therefore, we can assume that an invention that is not patented is not available to designers and does not directly contribute to the evolution of a technology area.

Future research will attempt to use patent maps analysis in order to try to predict future trajectories of technology development with the goal of technology forecasting.

REFERENCES

Campbell, R.S. (1983), Patent trends as a technological forecasting tool. World Patent Information, 5(1979), pp.137–143.

Cash, P., Stanković, T. & Štorga, M. (2013), Using visual information analysis to explore complex patterns in the activity of designers. Design Studies, 35(1), pp.1–28.

Dedehayir, O., Nokelainen, T. & Mäkinen, S.J. (2014), Disruptive innovations in complex product systems industries: A case study. *Journal of Engineering and Technology Management* - JET-M, 33, pp.174–192.

Kim, Y.G., Suh, J.H. & Park, S.C. (2008), Visualization of patent analysis for emerging technology. Expert Systems with Applications, 34(3), pp.1804–1812.

Lee, S., Yoon, B. & Park, Y. (2009), An approach to discovering new technology opportunities: Keyword-based patent map approach. Technovation, 29(6–7), pp.481–497.

OuYang, K. & Weng, C.S. (2011), A New Comprehensive Patent Analysis Approach for New Product Design in Mechanical Engineering. Technological Forecasting and Social Change, 78(7), pp.1183–1199. Available at: http://dx.doi.org/10.1016/j.techfore.2011.02.012

Smojver, V., Štorga, M. & Potočki, E. (2016), An Extended Methodology for the Assessment of Technical Invention. In S. Marjanović, Dorian ; Štorga, Mario ; Pavković, Neven ; Bojčetić, Nenad ; Škec, ed. P*roceedings of the 14th International Design Conference (DESIGN 2016)*. Design Society, pp. 1135–1144.

Smojver, V., Štorga, M. & Potočki, E. (2016), An extended methodology for the assessment of technical invention evolution. *Proceedings of International Design Conference*, DESIGN, DS 84, pp.1135–1144.

Son, C. et al. (2012), Development of a GTM-based patent map for identifying patent vacuums. Expert Systems with Applications, 39(3), pp.2489–2500. Available at: http://dx.doi.org/10.1016/j.eswa.2011.08.101

Su, H.N. & Lee, P.C. (2009), Dynamic and quantitative exploration on technology evolution mechanism: The case of electrical conducting polymer nanocomposite. PICMET: Portland International Center for Management of Engineering and Technology, Proceedings, pp.2433–2440.

Suh, J.H. & Park, S.C. (2009), Service-oriented Technology Roadmap (SoTRM) using patent map for R&D strategy of service industry. Expert Systems with Applications, 36(3 PART 2), pp.6754–6772. Available at: http://dx.doi.org/10.1016/j.eswa.2008.08.062

Suzuki, S. (2011), Introduction to Patent Map Analysis. Japan Patent Office.

Štorga, M., Mostashari, A. & Stankovic, T. (2013), Visualisation of the organisation knowledge structure evolution. Journal of Knowledge Management, 17(5), pp.724–740. Available at: http://www.emeraldinsight.com/journals.htm?issn=1367-3270&volume=17&issue=5&articleid=17097363&show=html

Tseng, Y.-H.H., Lin, C.-J.J. & Lin, Y.-I.I. (2007), Text mining techniques for patent analysis. Information Processing & Management, 43(5), pp.1216–1247. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0306457306002020

Yang, S.Y. & Soo, V.W. (2012), Extract conceptual graphs from plain texts in patent claims. Engineering Applications of Artificial Intelligence, 25(4), pp.874–887. Available at: http://dx.doi.org/10.1016/j.engappai.2011.11.006

Yoon, B.-U., Yoon, C.-B. & Park, Y.-T. (2002), On the development and application of a self-organizing feature map-based patent map. R and D Management, 32(4), pp.291–300. Available at: http://doi.wiley.com/10.1111/1467-9310.00261

Yoon, B. & Park, Y. (2004), A text-mining-based patent network: Analytical tool for high-technology trend. *Journal of High Technology Management Research*, 15(1), pp.37–50.