INDUSTRIAL CHALLENGES IN PATENT MANAGEMENT FOR DESIGN INNOVATION

Gokula Vasantha, Andrew Wodehouse, Jonathan Corney, Ross Maclachlan, Ananda Jagadeesan
Design Manufacture and Engineering Management, University of Strathclyde, Glasgow, UK
gokula.annamalai-vasantha@strath.ac.uk, andrew.wodehouse@strath.ac.uk, jonathan.corney@strath.ac.uk, r.maclachlan@strath.ac.uk, ananda.jagadeesan@strath.ac.uk

Abstract
Innovation is critical to sustain in prevailing competitive business environments. Industries need effective innovation strategies in-practice to develop and deliver novel products and services swiftly. In order to implement innovation strategies effectively, industries need innovation capacity in engineering design supported with intellectual assets. However, there are many issues that prevent streamlining these processes. The objectives of this research are to explicit the issues related to industrial patents (one of the important resources in intellectual assets) generation and management processes, and propose cost-effective crowdsourcing approach as a tool for patent landscaping activities. Interviews with patent attorneys and intellectual audit specialists reveal that most industries have ineffective intellectual property strategy; engineers do little patent searching, face challenges to identify novel product features, and often find difficulties to interpret patent information. The initial experiments of using the crowdsourcing approach for patent clustering activity reveal that general crowd workers (not knowing much about patents) were able to identify one third of expert clustered schema for much lesser cost. Further research work to strengthen the usefulness of the crowdsourcing approach for patent landscaping related activities is discussed.

INTRODUCTION
Continuous innovation is required for any business to keep pace with changing market requirements and adapt to challenges posed by global competitors. To grow continuous innovation culture in engineering design, industries should formulate strategies particularly focusing on understanding and utilizing their own intellectual assets appropriately (predominantly on patents), explore opportunity to buy leveraging patents, and conduct patent landscaping to watch closely on competitors' patents and intend to identify potential gaps for upcoming technologies (Trappey et al. 2012). A patent landscape analysis is defined as "a state-of-the-art patent search that provides graphic representations of information from search results" (Yang et al. 2010). Patent landscapes are used in R&D planning, technology positioning, design innovation and technology forecasting. Currently, patent landscaping is predominately undertaken by patent attorneys using software such as Thomson Themescape Maps™. This process is time consuming and costlier for smaller industries to undertake and also there is no easy process to evaluate outcomes generated (e.g. comprehensiveness of referred patents). In line with this observation, Bubela et al. (2013) argued that lack of transparency and inconsistencies persist in current landscaping practices. There is a need to help industries

develop quicker and less costly approach for patent landscaping focused to their field of engineering design, and to develop effective business innovation strategies.

This paper describes ongoing research to establish a strong empirical base on industrial issues surrounding current practices in patent processes, and to propose an alternative cost-effective and more responsive crowdsourcing approach for patent landscapes. The empirical base is generated by analyzing transcripts of expert patent attorneys' interviews. Interviews were structured in an informal format (rather than structured), providing opportunities for experts to immerse and reflect on processes and difficulties encountered in real-time patent processes. The transcripts were categorized to identify common themes (i.e. challenging processes) emerging between experts' viewpoints. In parallel to these interviews and data analysis process, an alternative mechanism to conduct patent landscaping is proposed through a crowdsourcing approach. There is a general assumption that manual extraction of desired patents is very costly and time consuming (Russo and Montecchi, 2011). However, using a crowd for various patent analysis activities could undermine this assumption. Proposing a crowdsourcing approach for patent analysis is not novel since companies like Article One™ operate in this domain. However, the key differentiation of this research is to develop an effective low cost solution through using a general crowd (i.e. not experts in patent searches) to create better patent landscaping outcomes.

As a first step towards crowd landscaping, a patent clustering task was experimented in the mTurk™ and CrowdFlower™ platforms. The crowdsourced results are compared with the patent clusters generated by computational algorithm proposed in an academic paper (Fu et al. 2013b). Empirical conclusions from interviewing expert patent attorneys combined with results from the initial crowdsourced cluster task are used to develop a case for advancing the crowdsourcing approach for patent landscapes. We additionally outline how this approach can eventually be formulated to help industries quickly and effectively to develop support tool for engineering design and identify business innovation strategies.

The following sections of the paper detail a literature summary of business innovation strategies, engineering design and patents, research questions and methodology, experts' views on challenges in generating and managing patents, initial results for patent clustering task with crowdsourcing approach, and finally conclude with discussion.

BUSINESS INNOVATION STRATEGIES, ENGINEERING DESIGN AND PATENT CLUSTERS – A LITERATURE REVIEW

Innovation strategy aims to propose an "innovative direction of business approach to the choice of objectives, methods and ways to fully utilize and develop the innovative potential of the business" (Hittmar et al. 2015). It intends to help structuring what, why and when to carry out innovation activities (Tidd et al. 2007). This strategy intends to help industries to develop new products, services or technological processes which eventually lead to improve business performances. Although innovation strategy is widely emphasized for business sustainability, Kubina (2011) argued that developing innovation strategy is a complex activity due to involvement of all business entities. Also innovation strategy varies based on vision and mission, and primary drivers of industries. Some of the classifications based on drivers and mission are tabulated in Table 1. Additional innovation strategy classifications are listed in Akman and Yilmaz (2008).

Table 1. Classification of innovation strategies based on business drivers and objectives

Source	Classification
pwc, 2014 (referred in Babkin et al. 2015)	Need seekers, market readers, technology drivers
Rodionov et al. 2014	Make and buy strategies
Hultink and Robben, 1995	Technological innovator, rapid copier, cost reducer
Ansoff and Stewart, 1967	First to market, follow the market leader, applied engineering, develop me-too products
Manu and Sriram, 1996	Product innovator, process innovator, late enterer, non-innovator, original initiators
Lynn and Mazzuca, 1998	Customer oriented, process oriented, initiator oriented, learning oriented
Lee et al. 2009	Complementation, Synergy, Exploration, Expansion

Analysis of these classifications reveals that customer, competitor, and technological focus are the key common and critical factors in developing any innovation strategy. This higher level innovative strategy is directly linked to the innovative capacity of every organization. A firm's innovative capacities depend on strategy, organization, learning, processes, and networks (Tidd and Bessant, 2009). One of the elements that interlink innovative strategy factors with innovative capacity in engineering design is intellectual assets, because they aid to compare competitors and technological growth with reference to internal developments. Figure 1 explains connectedness between innovation strategy, innovative capacity and intellectual assets.

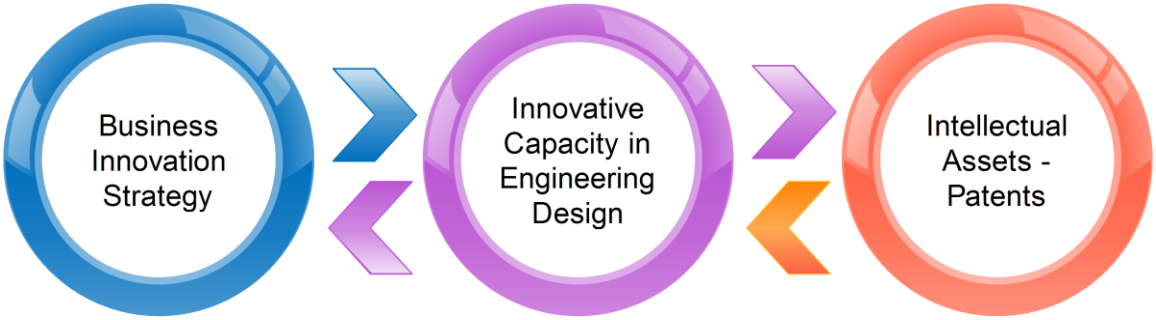


Figure 1. Connectedness between innovation strategy, innovative capacity and intellectual assets

Among many intellectual assets (e.g. registered designs, standards, trademarks, patents, technical know-how, trade secrets), patents are an important knowledge source that play a vital role in understanding competitors' focus and technology development. Kim and Lee (2015) argued that patents are regarded as important sources of innovation and are closely related to innovation strategies. Also patents act as an important knowledge source for employees' learning skills. There are various papers detailing patent related mechanisms that enable innovation and pro-competitive benefits (Kim et al. 2016). For example, patents help in assessing solution's novelty, identification of prior-art, business

gaps identification through patent landscapes, knowledge transfer, and technology forecasting. This paper focuses on identifying the practical industrial challenges in generating and managing patents, and experimenting usefulness of the crowdsourcing approach for patent clustering. Understanding patent-related industrial issues and crowdsourcing benefits will help to support engineering design processes and structure innovation strategies. The subsequent paragraphs review state-of-the-art in patent clustering focused to engineering design domain.

Patent clustering is a process of grouping related patents and representing graphically to support and enhance many patent-related applications. Patent clusters help to assess degree of importance, newness and similarity of patents (Yoon and Park, 2003). Clustering process has been undertaken at various levels of detail. For example, clustering is carried out at the document level (Smith, 2002), the structural information (e.g. patent claims) (Kim and Choi, 2007), the textual content of patents (Trappey et al. 2006), citation of patent pairs (Lai and Wu, 2005), patent network (Yoon and Park, 2003), and actor-relation map (Lee et al. 2009). Fu et al. (2013b) systematically evaluated a Bayesian-based computationally-generated patent clustering generated for 45 patents (Figure 2) with four expert designers' mental models of the domain. They argued that the computationally-generated clustering of patents and organization of clusters are sensible, intuitive, and find common ground with experts. They emphasized that this generated patent cluster could be used as an automated tool applied to design-by-analogy. In another related work, Fu et al. (2013a) compared patent clusters generated from functional and surface based mapping, and interestingly studied different forms of structure that could be used in patent clusters (e.g. ring, chain, tree etc.). Liang and Liu (2013) proposed a feature association-based approach to cluster patents to represent the categories of invention from a corporation perspective. This clustering approach is based on extraction of issue, solution and artifact information from patents. Using patents of Apple Inc. corporation, they demonstrated key artifact groups emerging from patent clusters. Since this clustering is based on particular extracted information, it is a challenge to design computational approaches for the multiple dimensions of information.

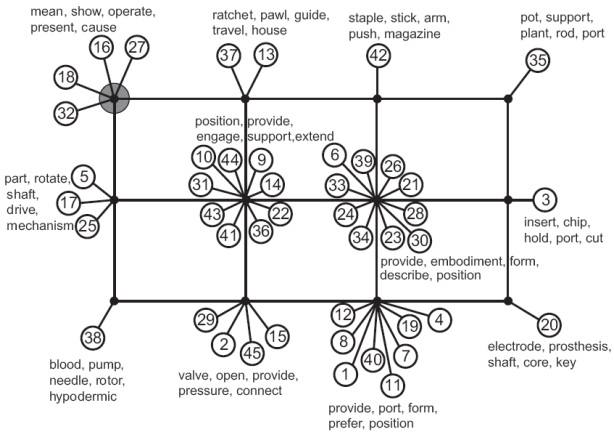


Figure 2. Computationally generated patent cluster and labels (Fu et al. 2013b)

Cong and Tong (2008) introduced an automatic patent classification approach for TRIZ users. They classified inventive principles based on descriptive information into obscure

and distinct inventive principles. Distinct inventive principles are further classified based on text and meaning similarity. They addressed multi-label classification issue and grouped patents based on the TRIZ Inventive Principles to help designers to search for patents based on solutions. Kitamura et al. (2004) sketched possibility of using patent map of functional decomposition to be used for the design review. Russo et al. (2012) proposed a patent search strategy based on a Function – Behavior – Physical effect – Structure ontology using natural language processing (NLP) techniques. They demonstrated the developed approach by building tree diagram for the state of the art of a nutcracker. The limitations in this approach are the identification of the main function is up to the user, uncertainty in user's behavior changes due to this intervention, and incompleteness in physical effects thesaurus.

Considering there are many approaches proposed for patent clustering, it is clear that there is not necessarily one best way to structure patent clusters. In other words, there is no correct solution for answering "how to best structure design repository data or patents". Identifying optimal structures for patent clustering is therefore ripe for novel solutions. Bubela et al. (2013) argued that lack of transparency and inconsistencies persist in current landscaping practices. Yoon and Park (2004) analyzed that the patent network may be ambiguous or meaningless if the structural relationship among patents is unclear. Although there are many approaches proposed in the literature, their real-time application in engineering design is questionable. Also commercially available patent landscaping software programs are expensive. The majority of available landscaping tools are primarily provided for high level strategy, market type decision making. Whilst there is some research in landscaping/clustering/dissecting patents for re-use in design, there is currently no tool optimized to leverage patents for design. There is a need to help industries develop quicker and less costly approaches for patent landscaping focused to their field of engineering design to develop effective innovation design and strategies.

In this research, we aim to conduct patent landscaping through a crowdsourcing approach using a general crowd who are not experts in patent searches. There are few research works which have used the crowdsourcing approach in engineering design. Vattam and Goel (2011) used an expert-based crowdsourcing approach to cataloguing and annotating research articles using the SBF-based approach to create, Biologue, a social citation cataloguing system. Wu et al. (2015) proposed a crowdsourced design (cDesign) framework to support generation and evaluation of crowd-enabled design activities. However, using the crowdsourcing approach for detail patent analysis in engineering design is not yet reported. The next section details the research questions and methodology used.

RESEARCH QUESTIONS AND METHODOLOGY

The research questions answered in this paper are:

1. What are the challenges faced by industries in generating and managing patents?
2. Could the crowdsourcing approach be effective to generate patent clusters with appropriate group labels?

The first question is answered by conducting interviews with five experts in patent processes. The experts represent an intellectual audit specialist (IAS), two patent attorneys (PA1 and PA2), an open innovation manager (OIM), and a patent informatics analyst (PIA). The abbreviations mentioned in the brackets are used as person identifiers

for informal interview transcripts quoted in the next section. Interviews are structured in an informal format (rather than structured), providing opportunities for experts to immerse and reflect on processes and difficulties encountered in real-time patent processes. The interviews occurred approximately for an hour each. The transcripts are categorized to identify common themes (i.e. challenging processes) emerging between experts' viewpoints.

To answer the second research question regarding patent clustering, we have used 45 patents selected by Fu et al. (2013b). This set of patents was selected because it aids to compare the crowdsourcing results with a computer-generated patent clustered structure (Figure 2). Most of these patents are classified within the U.S. Patent classification system as "Body treatment and care, heating and cooling, material handling and treatment, mechanical manufacturing, mechanical power, static, and related arts". In addition to patent clustering, the results analyzed also the group labels generated by computation algorithm, experts, and crowd workers. The crowdsourcing experiments were conducted in Crowdflower™ and MTurk™ crowdsourcing platforms. The subsequent sections structure the answers obtained for the above-mentioned research questions.

CHALLENGES IN GENERATING AND MANAGING PATENTS – EXPERTS' VIEWS

This section presents challenges mentioned by experts in utilizing and managing patents effectively for organizational benefits. The issues mentioned are discussed along with informal interview transcripts and related literature. The challenges are grouped and discussed in the following sub-sections: organizational problems, patent searches and landscapes, using patent information, and novelty assessment.

Organizational problems

Industries should have proper intellectual property (IP) management strategies to protect and benefits from them. However, "Do industries have proper IP strategies?" Unfortunately it looks like the answer is "no" for most of the industries. Our interviews revealed that:

"It takes up to 3 years to persuade a quite successful engineering company that intellectual property is something they should do." "In most cases there is no corporate division on IP." - IAS

Even if companies apply for patent grants, they struggle to get patents, and paybacks from granted patents. The UK Intellectual Property office observed that only about 30% of all the applications received by the IPO go on to become granted patents (Intellectual Property Office, 2012). In the worst case, they could lose currency due to inappropriate IP strategies for granted patents. All the scenarios mentioned by different interviewees are summarized.

"Not all applied patents are published, but most of them get published after year and half of filed date. Once they published they won't make it to grant. Some of them drop before grant." – PA1

"Companies struggle to keep track of what products are covered by what bits of patent applications, trademarks or other. They don't necessarily know what (patent) is still enforced, and what is expired. They don't keep track whether the new product is still covered by old patents they got. They don't know which countries they (patents) are covered in." – PA1

"We did go through the exercise of trying to cut out the majority of the patents that are not cited, but that have been active for 25 years and constantly paid for (renewal fee). Those patents are not worth protecting to be honest." – OIM

"Patent process is slow (up to 5 years to grant). So once the patent is filed, they don't keep track of it, they don't know what happened to it" – PA1

"We don't have good understanding of our own and competitors patents." "Don't have the idea of how many (patents) we have at this moment. IP practice is very much ad hoc (project specific)." – OIM

"There is going to be patents sitting down there, the core expertise we have but we don't really use. Potentially we can actually look (these patents) for licensing out." – OIM

Industries should manage patents in all IP stages starting from patent generation to leveraging benefits from them (i.e. looking for innovation, patent filed, published, granted and commercialized). To change these existing sub-optimal scenarios, the UK government is financially supporting SMEs to conduct IP audits (Intellectual Property Office, 2014). The evaluation of this supporting scheme revealed that 17% of the participated companies have filed patents, and 43% identified new opportunities to exploit their IP through initiatives like licensing and franchising. Although these results look encouraging, the companies participating in this scheme are still move slowly to implement recommendations due to costs and resource limitations. The UK government is also encouraging businesses to effectively use patents by providing a lower rate of Corporation Tax to profits earned from its patented inventions (HM Revenue & Customs, 2007). Since sustainability of these support schemes from the government could be limited, industries need to understand the benefits and find ways to solve the above mentioned issues.

Patent searches and landscapes

Searching patents plays a vital role in all patent related activities such as patent landscaping, product novelty assessment, and patent knowledge transfer. Although there are many computational techniques proposed in literature for patent searching (Bonino et al. 2010; Wang et al. 2015; Montecchi et al. 2013), keyword searches and patent classification filtering are the commonly used modes of searching. The interviews with experts revealed that industries find patent searching a difficult activity due to inaccessibility of different databases and patents in different languages. They often outsource these activities due to these limitations.

"It's quite difficult to search (believe it or not!!). That's why we don't search patents ourselves. It's a skilled job. The best you could do is search for keyword in title and abstract." – PA2

"There are few tools companies are looking at: how do you search all patent documents including (patents) in different languages using keywords and classification." – PA1

The keyword searches are not adequate to find appropriate documents due to inaccurate usage of terminologies, synonyms, polysemy, pronouns, multiple attributions, varying detail levels of patent descriptions, and homographs present in patents (Vasantha et al. 2016). Also the interviewees negated patent classification system due to difficulties in finding analogous solutions.

"Sometimes you want to solve a problem which is analogous to something but it could be in different classification codes, then patent searches may not bring these documents." – PA1

"Classification system is used by examiner to find things. If somebody comes with something new, it might have few classification marks, because there is nowhere to put it." – PIA

Apart from the issues mentioned in the search process, another important issue observed in the interviews is that innovators do little patent searching. This observation is also highlighted in the UK survey results where the "patent" category was noted only once within the whole 852 knowledge and information needs mentioned from 129 managers and engineers (Heisig et al. 2010). The major reasons cited in the interviews for poor patent searches are having good knowledge in the particular field, and most innovations are incremental. However, the implication of fewer searches is severe (e.g. filing patents for already existing ones).

"Inventors got reasonable idea about their field. Most of our work is for incremental innovation based on existing products. They have good idea of what competitors are up to. But they don't do much (patent) search." – PA1

"Most of the engineers don't think patents are part of their world." – IAS

"Lots of people who design products don't refer to patent literature. They think they know the area. Often they are wrong." – IAS

"You mentioned that two patents are exactly similar. But if you looked at patent applications you might get hundreds of patents that are identical. People will file patent not realizing some years before somebody already done it (exactly same thing). Sometimes it could be companies own prior art (they didn't bother to look into their own patent portfolio)." – PIA

Patent landscapes often incorporate flashy visuals which capture people attention by grouping related patents, and representing graphically to support and enhance many patent-related applications. Landscaping is an interactive tool through which people could compare their patent portfolio with their competitors. The current perception of landscapes is that there is not necessarily one best way to structure patent clusters.

"There is no recognized approve of artistic impression of how to do it (patent landscape)." – PA1

Using patent information

The issues not only stopped at the patent search process. Even if engineers get correct patent documents for their work, it is argued that interpretation of patent information is a

greater challenge. The challenges are many folds: patent information contains a complex blend of text and graphical content with legal entities, vagueness in graphical description (may be due to concept not adequately developed), patent information changes substantially in the patent grant process, variation in patenting process across the globe and language issues, and intention to avoid detailed information to make broader claims. It is argued that the patent is not product leverage (i.e. mapping patents to products are difficult). Surden (2011) provided an overview of various issues due to uncertainty in patent interpretation such as cost increase, decrease innovation, and scope underestimation.

"Interpretation of patent claims to picture could be complex, and patent without image is more difficult. Without picture and (description) numbers, you will get huge variations in how they are interpreting it (patent claim)." – PA1

"The original patent application will change quite the few times before you actually get a patent. When you ask to interpret the patent it doesn't mean that what is actually granted and that's what link to successful product." – PIA

"You will get lots of line drawings, as oppose to what goes on the product. During patent filing lots of people don't have final design yet. And also equally they don't want to give big attention to it because it buys extra secrecy even if it gets published. It still line drawings with couple of boxes, and flow diagrams. Nobody can work out other than the designers (inventors)." – PIA

"You really want to extend your monopoly as much as possible. That's why patents are still granted in words, and you interpret those words in relation to what has been filled. Picture could be one embodiment of your invention, it could have another embodiment. If you read the claim, it could be constructed differently." – PIA

Novelty assessment

Hall et al. (2013) observed that only about 4% of the firms in the UK are patenting among those who reported that they have innovated. Our interviews with experts reveal the possible reasons for this scenario could be in identifying and understanding what could be patentable. All the below transcripts reiterated these two factors repeatedly.

"Often people will have the wrong impression of where there unique elements are." – IAS

"It takes two years to pursue that it could be patented. They said they just put together some existing bits, there is nothing novel we brings into this product." – IAS

"In some cases engineers are resistant that they have done anything inventive. They have elevated idea of what could be patentable" – IAS

"We need to work to tell you that what you should be looking for something novel and then to fix for the product and then attach means to the patent law scope." – PA2

"So if you are sitting and thinking, if I could use that thing and combined this piece of information that is actually inventing. You have actually invented something. If you work day-in and day-out in the R&D environment you will just forget that. We had that problem all the time with the clients." – PA2

Interviews summary

Figure 3 summarizes the issues discussed in the above paragraphs with important interview transcripts. Although these analyses involve only 5 participated experts, the cross-sector profile and in-depth nature of interviews in relation to the practical issues faced by innovators we believe highlight important issues faced by industries in effective use of intellectual property.

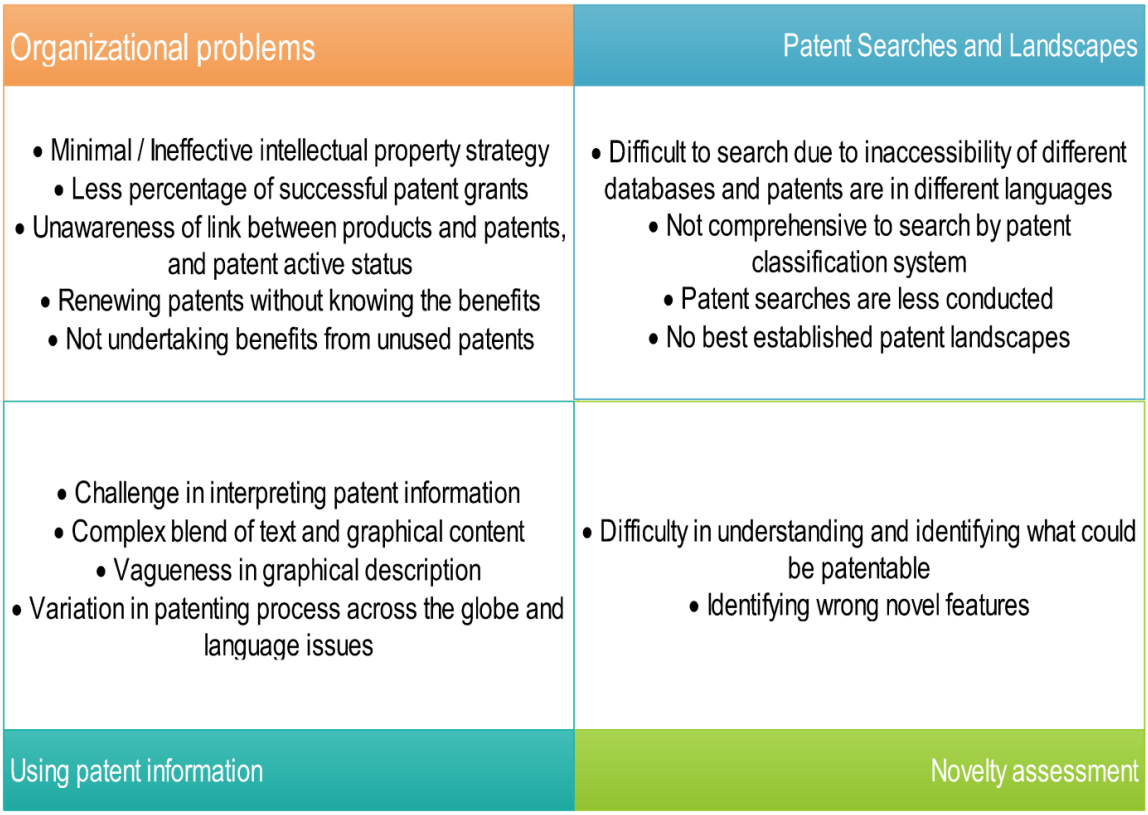
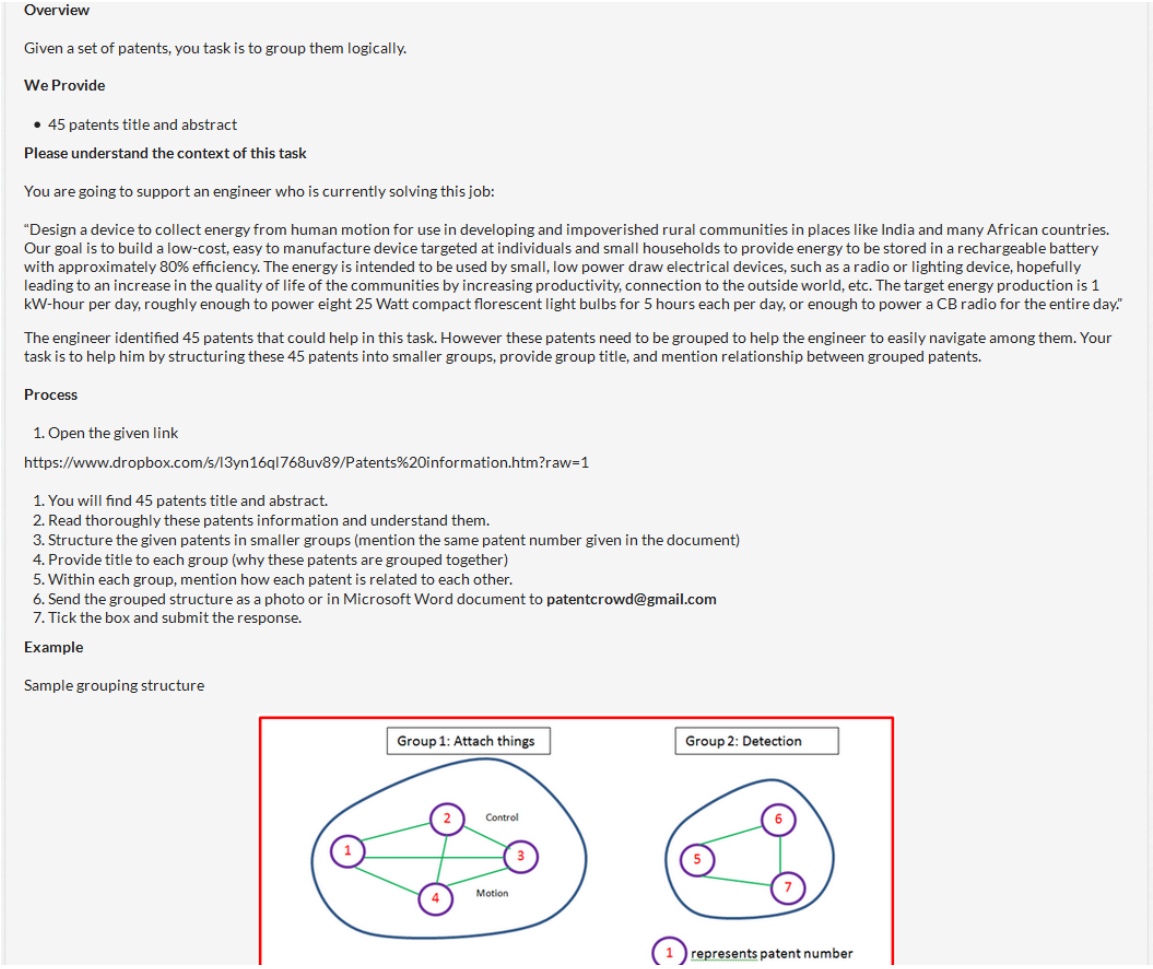


Figure 3. Summary of patent related issues identified from expert interviews

CROWDSOURCING PATENT CLUSTERING TASK – INITIAL RESULTS

The aim of this research is to test whether the crowdsourcing approach could be an alternative mechanism for patent analysis which cost less and brings in additional benefits to organizations in comparison to software programs. In this initial experiment, we reported the crowdsourcing results generated for a patent clustering task. We framed this task in-alignment with the exercise conducted by Fu et al. (2013b). Fu et al. (2013b) used 45 patents to understand how an expert might organize these patents to facilitate searching for potentially relevant inspiration for design-by-analogy. They compared experts clustering to a Bayesian-based computational algorithm cluster. Our aim is to compare these results (i.e. both from experts and computational clusters) with crowd generated clusters. This comparison should help to understand commonalities and differences among these approaches. Fu et al. emphasized in their laboratory experiments that experts should focus on 'functional similarity' (i.e. what the object does) in clustering patents. In this research we did not provided this focus because it will be added burden on the crowd who are not familiar with this kind of task and patents. We used

MTurk™ and Crowdflower™ crowdsourcing platforms to conduct the patent clustering task. The task posted on these platforms is illustrated in Figure 4. We paid on the average of $1.5 to the participated crowd for successful completion of this task.



Overview

Given a set of patents, you task is to group them logically.

We Provide

- 45 patents title and abstract

Please understand the context of this task

You are going to support an engineer who is currently solving this job:

"Design a device to collect energy from human motion for use in developing and impoverished rural communities in places like India and many African countries. Our goal is to build a low-cost, easy to manufacture device targeted at individuals and small households to provide energy to be stored in a rechargeable battery with approximately 80% efficiency. The energy is intended to be used by small, low power draw electrical devices, such as a radio or lighting device, hopefully leading to an increase in the quality of life of the communities by increasing productivity, connection to the outside world, etc. The target energy production is 1 kW-hour per day, roughly enough to power eight 25 Watt compact florescent light bulbs for 5 hours each per day, or enough to power a CB radio for the entire day."

The engineer identified 45 patents that could help in this task. However these patents need to be grouped to help the engineer to easily navigate among them. Your task is to help him by structuring these 45 patents into smaller groups, provide group title, and mention relationship between grouped patents.

Process

1. Open the given link

https://www.dropbox.com/s/l3yn16ql768uv89/Patents%20information.htm?raw=1

1. You will find 45 patents title and abstract.
2. Read thoroughly these patents information and understand them.
3. Structure the given patents in smaller groups (mention the same patent number given in the document)
4. Provide title to each group (why these patents are grouped together)
5. Within each group, mention how each patent is related to each other.
6. Send the grouped structure as a photo or in Microsoft Word document to **patentcrowd@gmail.com**
7. Tick the box and submit the response.

Example

Sample grouping structure

Figure 4. Patent clustering task illustration to the crowd

The crowd responses from these platforms were collected in batches. Without any quality control in place, the percentage of valid crowd responses received from MTurk™ platform (80%) is much higher than Crowdflower™ platform (10%). This high variation in receiving appropriate responses between platforms illustrates that the crowd task has to be customized for each and every crowd platform, and quality control steps should be incorporated during the crowd experiments to increase valid responses. One of the best clustering generated from the crowd is illustrated in Figure 5. Table 2 compared time taken among crowd, expert and computational algorithm. The comparison reveals that all the three approaches take approximately 1 hour to complete this patent clustering task. The crowd approach could take more hours to complete, if the number of responses requested (i.e. number of people to respond) increased. In the presented case, the posted task was completed in two days because the number of responses requested was 100. In total, we received 18 valid crowd patent cluster responses from both tested platforms. These clusters are analyzed and presented in the subsequent paragraphs.

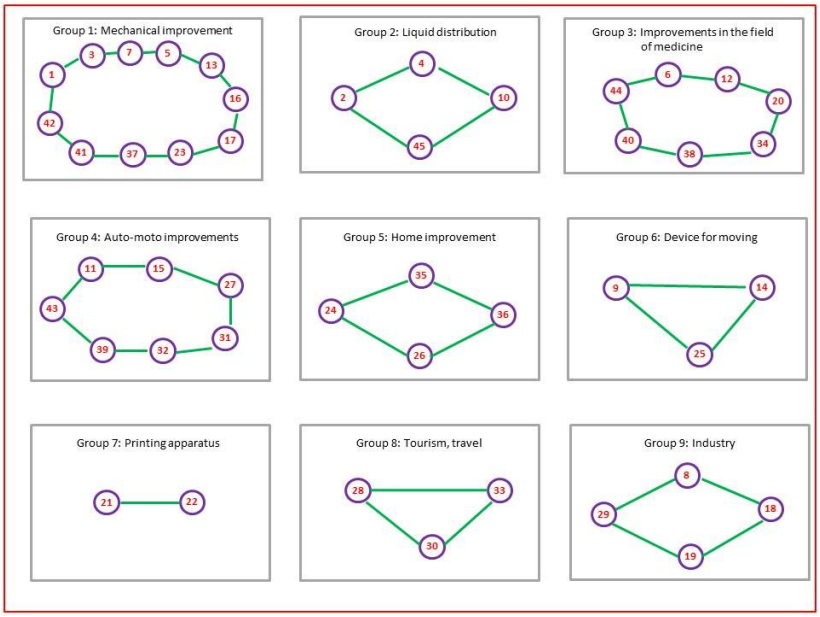


Figure 5. Sample patent cluster groups generated by a participant from crowdsourcing platform

Table 2. Time taken and average number of patent clusters from the three approaches

	Crowd	**Experts**	**Computational platform**
Time taken	Less than 1 hour to 2 days	Approximately 1.25 h (Fu et al. 2013)	Less than 45 minutes to optimally generate 8 different structures (Fu et al. 2013)
Average number of clusters	8	8	12

Comparison of the average number of cluster (Table 2) reveals that both the crowd and expert generated similar cluster size, but less than the computational algorithm outcome. To study equivalent of these clusters, cluster labels are studied. Since the labels generated from the computational algorithm are based on the top five highest average ranked words for that cluster (Figure 2), unlike experts and crowd marking with a single label, these computational generated cluster words are not taken into account in our analysis.

Table 3. Comparison between experts and crowd clusters

	Experts	**Crowd**
Number of different clusters generated	35 clusters from four experts	57 clusters from 18 participated crowd
Number of common clusters	11	

Table 3 compares different variables between experts and crowd. The analysis reveals that cluster convergence among experts is very minimal (11%) compared to crowd (40%). On average, the four experts generated 8 different cluster labels (i.e. almost everything is different among themselves), whereas crowd generated only 3 different clusters labels/person. The interesting finding is that the crowd could able to find one third of experts' clusters (31%). Table 4 lists the similar clusters identified between experts and crowd. This good similarity percentage is a good starting point to encourage using the crowdsourcing approach for patent clustering task. Experts used 'not interesting' cluster to group patents that are not relevant to the given design problem. This expert behavior is observable with two participants from the crowd. They categorized irrelevant patents in 'others' and 'left field' categories.

Table 4. Common clusters identified between experts and crowd

Common ten groups identified between experts and crowd	
Adaptable Better adaptability	Measuring and detectors, gauges collection of information. Measurement Devices & Sensors
Add on for travel Tourism, Travel	Solar and wind power Floor and solar work
Attachments Attach things to the body Attach/connect to cart Mechanical joint Combining elements Mechanical linkage Fastener/ office Mechanical Construction Mechanical Fasten and control	Hydraulics Fluid dispensing and values Values/pump Hydraulics motors Fluids. Valves. And Fluids life supports. Hydraulics motors. Valves & Regulation Systems Fluid and fluid flow Liquid distribution Fluid delivery and regulation method
Motion Collect power from mechanical motion Discrete motion Device for moving Collect power from cart	Mechanical Feeders/Sorters Probe guide and dispensary method (ways to feed and dispense) Mortars / sorting Loads handling and moving Conveyers Sort and control
Frame mount Frames, Structures & Stands	Mix both Fluid mixing process

CONCLUSIONS

Interviewing experts in patent related fields reveal that industries are not leveraging extensive knowledge sources available in the form of patents. Industries have issues in all the stages of patent processing: identifying patentable concepts and features; applying, follow-up and getting patent grants; interpreting existing patents to identify potentials gaps for product development; and less patent searches leading to ineffective use of

patent knowledge sources. One of the possible reasons for these issues could be that industries do not have cost-effective support tools to aid in these processes. Most patent related activities are outsourced, considering these activities are time consuming and require specialist inputs. In this research, we aim to demonstrate a cost-effective crowdsourcing approach for patent analysis activities.

The initial experiments for patent clustering task demonstrate that a general crowd (who has no previous experience on patent processing) could able to generate one third of patent cluster schema produced by experts. This encouraging result demonstrates that with subsequent training providing to crowd could significantly increase crowd performances. It has been reported that experts were paid $100 for the 2h patent clustering study (Fu et al. 2013b). But the crowd participated in our initial experiments were happy to participate for receive less than $2. In view of computational software program results are not always sensible, this cost advantage could be greatly exploited by industries for various patent-related activities to develop support tool for engineering design, and sustainable business innovative strategies.

Although the crowdsourcing approach seems to be a viable option, there are challenges involved in getting useful results from the crowd. Ongoing research work is investigating approaches for incorporating appropriate quality control measures, motivational factors to encourage workers, and creating attractive and clear task instructions. Current work aims to evaluate the clusters generated by the crowd with experts to check potential 'fit for purpose' to design-by-analogy method. In expansion, the crowdsourcing approach will be further studied in the following patent-related design tasks: novelty check, technology survey, patent landscapes, identifying types of innovation (incremental, intermediate, or radical), and patent drawing interpretations. We believe that testing and applying crowdsourcing approach to the above mentioned tasks should help industries develop potential innovative products and strategies.

REFERENCES

Akman, G., & Yilmaz, C. (2008). Innovative capability, innovation strategy and market orientation: an empirical analysis in Turkish software industry. International Journal of Innovation Management, 12(01), 69-111.

Ansoff, HI and JM Stewart (1967). Strategies for a technology-based business. Harward Business Review, 45, (November–December), 71–83.

Babkin, A. V., Lipatnikov, V. S., & Muraveva, S. V. (2015). Assessing the Impact of Innovation Strategies and R&D Costs on the Performance of IT Companies. Procedia-Social and Behavioral Sciences, 207, 749-758.

Bonino, D., A. Ciaramella, and F. Corno, Review of the state-of-the-art in patent information and forthcoming evolutions in intelligent patent informatics. World Patent Information, 2010. 32(1): p. 30-38.

Bubela, T., Gold, E. R., Graff, G. D., Cahoy, D. R., Nicol, D., & Castle, D. (2013). Patent landscaping for life sciences innovation: Toward consistent and transparent practices. Nature biotechnology, 31(3), 202-206.

Cong, H., & Tong, L. H. (2008). Grouping of TRIZ Inventive Principles to facilitate automatic patent classification. Expert Systems with Applications, 34(1), 788-795.

Cong, H., & Tong, L. H. (2008). Grouping of TRIZ Inventive Principles to facilitate automatic patent classification. Expert Systems with Applications, 34(1), 788-795.

Fu, K., Cagan, J., Kotovsky, K., & Wood, K. (2013a). Discovering structure in design databases through functional and surface based mapping. Journal of mechanical Design, 135(3), 031006.
Fu, K., Chan, J., Schunn, C., Cagan, J., & Kotovsky, K. (2013b). Expert representation of design repository space: A comparison to and validation of algorithmic output. Design Studies, 34(6), 729-762.
Hall, B. H., Helmers, C., Rogers, M., & Sena, V. (2013). The importance (or not) of patents to UK firms. Oxford Economic Papers, gpt012.
Heisig, P., et al., Exploring knowledge and information needs in engineering from the past and for the future – results from a survey. Design Studies, 2010. 31(5): p. 499-532.
Hittmar, S., Varmus, M., & Lendel, V. (2015). Proposal of Evaluation System for Successful Application of Innovation Strategy through a Set of Indicators. Procedia Economics and Finance, 26, 17-22.
HM Revenue & Customs, 2007. Corporation Tax: the Patent Box, https://www.gov.uk/guidance/corporation-tax-the-patent-box
Hultink, EJ and HSJ Robben (1995). Measuring new product success: The difference that time perspective makes. Journal of Product Innovation Management, 12, 392–405.
Intellectual Property Office, 2012. PEERTOPATENT (UK) – PILOT, www.ipo.gov.uk.
Intellectual Property Office, 2014. IP Audits Evaluation Report, Final V3.0, Ref: JN1596/DD/JM.
Kim, B., Kim, E., Miller, D. J., & Mahoney, J. T. (2016). The impact of the timing of patents on innovation performance. Research Policy, 45(4), 914-928.
Kim, J. H., & Choi, K. S. (2007). Patent document categorization based on semantic structural information. Information processing & management, 43(5), 1200-1215.
Kim, J., & Lee, S. (2015). Patent databases for innovation studies: A comparative analysis of USPTO, EPO, JPO and KIPO. Technological Forecasting and Social Change, 92, 332-345.
Kitamura, Y., Kashiwase, M., Fuse, M., & Mizoguchi, R. (2004). Deployment of an ontological framework of functional design knowledge. Advanced Engineering Informatics, 18(2), 115-127.
Kubina, M. (2011). Realizacia podnikovych procesov pomocou BPM. In Modelovani, simulace a optimalizace podnikovych procesu v praxi: sbornik z konference, Praha: CSOP. pp. 223-230.
Lai, K. K., & Wu, S. J. (2005). Using the patent co-citation approach to establish a new patent classification system. Information Processing & Management, 41(2), 313-330.
Lee, S., Yoon, B., Lee, C., & Park, J. (2009). Business planning based on technological capabilities: Patent analysis for technology-driven roadmapping. Technological Forecasting and Social Change, 76(6), 769-786.
Liang, Y., & Liu, Y. (2013, August). Rationale-based patent analysis for corporate product design. In ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. V02BT02A013-V02BT02A013). American Society of Mechanical Engineers.
Lynn, GS and M Mazzuca (1998). Learning the critical success factor in developing truly new products. Research Technology Management, 41(3), (May/June), 45–53.
Manu, FA and V Sriram (1996). Innovation, marketing strategy, environment and performance. Journal of Business Research, 35, 79–91.

pwc, Global Innovation 1000 study microsite [electronic resource]. URL: http://www.strategyand.pwc.com/global/home/what-we-think/multimedia/video/mm-video_display/improving-innovation (Accessed: 25.02.2016).
Rodionov, D.G., Rudskaia, I.A., Kushneva, O.A. (2014). The importance of the university world rankings in the context of globalization, Life Science Journal, 11 (10), 442-446.
Russo, D., & Montecchi, T. (2011, January). A function-behaviour oriented search for patent digging. In ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 1111-1120). American Society of Mechanical Engineers.
Russo, D., Montecchi, T., & Liu, Y. (2012, August). Functional-based search for patent technology transfer. In ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. 529-539). American Society of Mechanical Engineers.
Smith, H. (2002). Automation of patent classification. World Patent Information, 24(4), 269-271.
Surden, H. (2011). Efficient Uncertainty in Patent Interpretation. Washington and Lee Law Review, 68, 1737.
Tidd, J., & Bessant, J. (2009). Managing innovation: Integrating technological, market, and organizational change. Chichester, England: John Wiley & Sons.
Tidd, J., & Bessant, J., & Pavitt, K. (2007). Řizeni inovaci. Zavadeni technologickych, trznich a organizacnich zmen. Brno: Computer Press.
Trappey, A. J., Hsu, F. C., Trappey, C. V., & Lin, C. I. (2006). Development of a patent document classification and search platform using a back-propagation network. Expert Systems with Applications, 31(4), 755-765.
Trappey, A. J., Trappey, C. V., Wu, C. Y., & Lin, C. W. (2012). A patent quality analysis for innovative technology and product development. Advanced Engineering Informatics, 26(1), 26-34.
Vasantha, G., Corney, J., Maclachlan, R., and Wodehouse, A. (2016). Use of crowdsourcing for the interrogation and presentation of patents in engineering design. Design Computing and Cognition conference, Chicago, USA.
Vattam, S. S., & Goel, A. K. (2011, June). Semantically annotating research articles for interdisciplinary design. In Proceedings of the sixth international conference on Knowledge capture (pp. 165-166). ACM.
Wang, J., W.F. Lu, and H.T. Loh, A two-level parser for patent claim parsing. Advanced Engineering Informatics, 2015. 29(3): p. 431-439.
Wu, H., Corney, J., & Grant, M. (2015). An evaluation methodology for crowdsourced design. Advanced Engineering Informatics, 29(4), 775-786.
Yang, Y. Y., Akers, L., Yang, C. B., Klose, T., & Pavlek, S. (2010). Enhancing patent landscape analysis with visualization output. World Patent Information, 32(3), 203-220.
Yoon, B., & Park, Y. (2004). A text-mining-based patent network: Analytical tool for high-technology trend. The Journal of High Technology Management Research, 15(1), 37-50.