Exploring the nanotechnology landscape for competitive advantage using SAO-mining

Roel C. Boekel, Delft University of Technology

ABSTRACT

Historically, nanotechnology faces the same challenges that every new, emerging and science-based industry faces. The transfer of viable knowledge from science to market needs to be bridged better. To enhance this transfer, a methodology is designed which integrates engineering design and subject-action-object (SAO) based text mining to find, select and evaluate nanotechnologies from scientific abstracts. A case study has been conducted for which the engineering design methods are used to structure a technical innovation problem in the manufacturing industry. 1.2 million abstracts from nanotechnology related articles are downloaded from the Web of Science and indexed based on technical function using SAO-parsing of the title and abstract. To this dataset, the methodology was applied to scout technologies solving a technical problem in the manufacturing industry. The performance of the designed methodology was measured based on precision and recall, as compared with standard scientific search retrieval technology. The methodology was evaluated on basis of structured interviews conducted with two engineering managers. The results demonstrate that the SAO methodology is a valuable assist to innovation and design within the firm. Significant challenges which are addressed in this article are for example knowledge loss due to SAO-parsing and information stickiness.

***Keywords*— SAO, Text mining, technology scouting, innovation, nanotechnology.**

I. Introduction

Nanotechnology is a tremendous, ground breaking field of science that enable materials to behave like never seen before [1], [2]. Special features at the nanoscale were the start of a research field where almost every natural science field meets. Knowledge of nanoparticle behavior, materials, surfaces, molecular motors and sensors resulted in a large crossover of knowledge between research fields that would never have been brought together [3]. Nanotechnology, is regarded as a general purpose and multi-purpose technology allowing that a single technology can be embedded and applied in many products, processes or systems [4], [5]. Until now the uncertainty of this ground-breaking technologies application areas, health aspects, industry organization and commercialization complexity have slowed the adoption of nanotechnology down [6]–[8]. This is very counterintuitive when reasoned from the perspective of technology driven companies. These technology driven companies have an ongoing need to keep innovating products and processes for short and long term survival [9]–[11]. It should therefore be more logical that companies embrace the developments in nanotechnology and actively adopt the enabling concepts in their products and processes.

To overcome the barrier of knowledge-transfer in nanotechnology, the complexity of nanotechnology commercialization and the nanotechnology industry needs to be bridged. On the one hand this means that companies must be enabled to evaluate nanotechnologies on its potential for innovation. On the other hand, this means that companies need to be able to find and select nanotechnologies. Nanotechnology is a science driven field of which the main challenge is how companies can understand and interpret science based knowledge. A practice that is regarded in business as complex, costly and inefficient [12]. Therefore, we have designed a methodology with a nanotechnology knowledge base which is operable by companies to find, select and interpret nanotechnologies for innovation problem solving. In the methodology we combine innovation theory, engineering design methodology and Natural Language Processing (NLP) techniques. Engineering design is used to structure a technical innovation problem within technology driven firms. Subject-Action-Object (SAO) based text mining techniques are used to organize nanotechnologies from 1.2 million nanotechnology related scientific abstracts from the Web of Science. By means of SAO-parsing the abstracts are indexed based on function[1]. The structured technical innovation problems are used as function-based input for browsing the science based nanotechnology toolbox. This search method enables linking functions described as Action-Object (AO) to Subject (S) which act as a means[2]. Allowing the user to find technologies and technological concepts based on the functions that need to be executed for solving the technical innovation problem. Furthermore, this novel technique builds further on the existing literature in patent based SAO-text mining and uses a different viewpoint for applying SAO by indexing scientific abstracts for finding, selecting and interpreting specific nanotechnological concepts.

This paper will discuss a case study in which the methodology will be applied, as developed for the manufacturing industry. In this case study we will carefully look how the methodology can be applied, if nanotechnology results are generated and to determine how well SAO based text mining information retrieval system performs to find science based technologies. This case study will examine if the methodology can be applied, whether nanotechnology results are generated and to determine how well SAO based text mining information retrieval system performs to find science based technologies.

The paper will be organized as follows. We will look into the theory and concepts about nanotechnology, innovation, technology scouting, engineering design and text mining. Based on these theories a methodology will be constructed in section III. In section IV, the case study will be introduced and presented. In section V the methodology will be thoroughly validated by means of internal verification, external validation and in depth interviews with R&D managers to reflect on the results. The pitfalls and challenges of the methodology are discussed and conclusions are drawn based on the results of the case study and performance of the methodology.

II. Theoretical framework

A. *Nanotechnology and technology transfer*

Nanotechnology is related to chemistry, physics and biology. Due to its small size, i.e. nano-scale, the behavior of the materials can be described by the laws of quantum physics [13]. This behavior allows the design of materials with tunable properties like higher surface to volume ratio, conductivity and semi-conductivity of materials, increased hardness, better chemical resistivity and improved transport kinetics [14]. For this reason the field of nanotechnology has gained the title of being a general purpose technology [15]–[17]. Therefore nanotechnology does not serve a single application but enables new possibilities in all kinds of industries. Against expectations, nanotechnology remains within science and is mainly adopted in high tech and science

[1] Functions are minimalistic (verb-noun) descriptions of what a certain system needs to do.

[2] Means are technologies or components that execute the function.

driven industries [18]. A gap has emerged in nanotechnology transfer from science to other technology driven industries than high-tech industries. Multiple theories exist why this gap has emerged within those industries, three main barriers are given; the diverged and fragmented industry [7], risks of nanotechnology in health and safety [19]–[21] and the gap for transferring science based knowledge to the market [6]. To bridge the gap between nanotechnology knowledge and mass adoption, uncertainty in these barriers needs to be reduced.

B. Innovation strategy in Bulk manufacturing

The need of innovation is indisputable for society to foster economical growth and for companies to secure firm survival in the long and short run [22]–[24]. Innovation is driven by technological development, increasing needs of customers, shorter product life cycles and the increased world competition [25]. Pavitt has argued that there are roughly three types of technology driven companies [26]: Supplier dominated, production intensive and science driven. Because supplier dominated firms heavily depend on input from suppliers and science driven firms already have close ties with knowledge institutes the interest is on the adoption of nanotechnology by production intensive firms. Within production intensive firms there are specialized firms and bulk manufacturing companies. Specialized firms mainly use design and development activities as a technology source while scale intensive companies also have their own production engineering department and R&D activities. Our main interest will be pointed towards these scale intensive companies that actively execute research and development activities. Companies that are technology or product driven, have the need to innovate, have an organized innovation structure, are unaware of the innovation potential of nanotechnology and actively sources for technology.

C. Innovation management, technology scouting and engineering problem solving in bulk manufacturing.

With knowing why companies innovate, the next task is to examine how firms organize their innovation. Many models have been proposed throughout time. The most well known concepts are the technology push and market pull model. In these concepts innovation is driven by technological development or a market need. These linear concepts have been discussed by Chesbrough and Kline and Rosenberg, who argue that that the innovation process is more iterative and organic than linear [27], [28]. These models also better align with the activities scale intensive companies undertake within their technology sourcing activities. This process is in literature described as the technology scouting process and covers the activities of identifying, evaluating and selecting technologies that suits the goals technical and challenges of the company [12], [29]. The required information can be found using multiple sources. These sources can be clients, competitors, universities, suppliers, internal knowledge, consultants conferences, exhibitions, on-line databases and patents [30]–[32]. The company organizes the technology scouting process as part of the innovation process and determines which knowledge can be applied in which situation. The tasks executed in the technology scouting process has overlap with the innovation management process of new product design. Rohrbeck has pointed out that both processes are subject to a phase of product specification and development [33], [34]. In this phase, requirements are constructed and the problems and goals are structured. In systems engineering this phase is essential for organizing the problem and to determine how a problem can be solved or a product can be developed.

D. Applying Text mining in technology scouting.

Due to ongoing growth of technical information, knowledge, and their accessibility it becomes more and more complicated for technology scouts within the companies to follow the latest trends and to track down the required information to solve problems [35]. Large amounts of this information is textually stored in databases and those databases can be searched to find valuable concepts in the innovation process. The ability to browse and structure text has led to the development of Natural language processing (NLP). A field in science that is able to train computers to interpret and search text like a human being [36]. Within technology scouting this technology is already broadly applied to find valuable information in technology related patents. Technology scouts use text-mining techniques to identify technology trends, construct technology trees, fill in technology roadmaps and for technology novelty detection [37], [38]. Within text mining many different sorts of techniques exist that allows to extract all kinds of information from text. Reviews by Abbas an Mattas have enumerated text mining techniques for patent analysis [37], [38]. Within these approaches five text mining approaches are discussed. NLP, function based, rule based, semantic based and neural network approaches. Of those techniques the NLP approaches are the most applied technique for patent analyses. Moreover, also function based approaches and semantic approaches are related to NLP. Two types of approaches can be distinguished within the NLP techniques itself; keyword based and syntax based SAO techniques. While keyword based approaches only have the ability to determine word frequencies and word similarities SAO-relations allow to extract relations between words in sentences. Consisting of a noun-verb-noun, these relations have the valuable possession of describing a problem-solution or input-output relation [39]. This functionality can be used in technical texts to determine functional components. The SAO-based approach is more powerful in extracting technical concepts from texts than keyword based approaches. Within patent analysis, apart from functionality extraction, SAO can be used for tree structuring of technology, technology roadmapping, trend analysis and novelty detection. It is interesting to further explore the application of SAO-mining in other technical texts than patents and to investigate how SAO-relations can be rendered for competitive advantage in nanotechnology.

III. Methodology

We have seen that companies in bulk manufacturing are in a continuous process of innovation of which technology is one of the main sources for realizing this. We have also seen that nanotechnology is a valuable multi-purpose technology toolbox that is not able to reach the market well enough. Through technology scouting practices, innovation managers should be enabled to integrate nanotechnologies in a structured way. However, the interpretation of scientific information is a complex and time consuming process and therefore this information should be presented in a manner that innovation managers are able to find, select and interpret the concepts efficiently. SAO-mining allows to extract problem-solution relations from text and therefore could be a valuable new approach for rapidly interpreting science based information. In the methodology we will integrate these theories and concepts.

A. Methodology requirements

Before designing the methodology, we first introduce requirements on what the methodology ideally should accomplish. The requirements relate to the internal and external validity of the methodology and to the performance of the methodology when applied. Essentially the requirements need to make sure that the methodology can be measured and assessed on its performance after application. For the SAO-based software part of the methodology this will mean that we will look at the internal verification and the external validity of the methodology [40], [41]. We will also look at the methodology from the viewpoint of the scale-intensive firm by means of semi-structured interview with the R&D manager. The main interest of the firm is to find and evaluate nanotechnologies for innovation purposes. But while doing so, the firm wants to increase the quality of the knowledge retrieved and reduction of cost of the information retrieval process. This way we can measure if the methodology and SAO based information retrieval system functions better than conventional information retrieval systems. Validity and methodology performance will be discussed in section V.

B. Methodology construction

The methodology consists of three parts. The first part aims at helping the firm to organize the innovation problem and converts this into input for the information retrieval system. The second part is the design of a nanotechnology innovation retrieval system which converts the input into interpretable nanotechnology knowledge concepts based on function. The third part is an evaluation step which will interpret the technology and measure which of the scouted technologies scores the best on the requirements determined in part one. The following subsections will describe the three steps in more detail.

1) Innovation problem solving

We have seen that the process of technical problem solving, in an innovation problem or an engineering problem is not a linear process. To solve a problem, it must be properly defined and structured. Engineering design methodology will be applied to structure the problem [33], [42]. This consists of defining the problem definition, designing case-requirements and functions of what the system needs to do. These structured problems then need to be answered using the SAO-based information retrieval system. Based on the problem definition measurable requirements are generated consisting of objectives and constraints. These requirements can be converted into functions of what the system requires to do. Within engineering design, functions can be described as minimalistic functions that are constructed of a verb and a noun [43]. In this relation the verb demonstrates the action and the noun the object. Good construction of functions will determine all the actions the solution should do. The minimalistic function also renders the Action-Object relations that can be used as input for the search model that we will describe in the next section.

2) Construction of the search model

For constructing a search model consisting of functional descriptions of nanotechnologies scientific abstracts are used as a data source. Nanotechnology is a science driven field and therefore extensively described in the scientific literature. To get the bulk of the nanotechnology concepts in a database, 1.2 million nanotechnology related article-abstracts are downloaded from the Web of Science. For every sentence in the title and abstract an algorithm is used to extract the SAO-relation from the sentence. The SAO-relation is tagged with a unique record number making it able to trace back the complete record (title and abstract). This database ordering enables to iterate through all the SAO-relations at high speed. The SAO-relations are cleaned by removing subjects, actions or objects that will not describe a technical relation. All the remaining SAO-relations are organized and stored in a single file. To browse the SAO relations a user interface is constructed with input-fields for the verb-noun functions constructed in step 1. To reduce the dependence on the semantic formulation of the function a synonym generator is embedded in the interface that allows to add words with same semantic meaning as the action or the object. The search model is able to iterate through the list with SAO-relations based on Action and Object. If a match is found in an Action and Object combination the SAO-relation will return to the screen. In the SAO-relations the AO-part of the relation shows the function which is executed by the subject (S) that acts as the means. The found means can then be further interpreted to show if the means can be a potential solution to (part) of the innovation problem. In the next section this interpretation will be further described.

3) Interpretation of results

For interpreting the results, we will use the requirements obtained in step 1. These requirements act as a measure to obtain if the found means can be potentially applied to solve the innovation problem. First all the subjects are selected. Subjects that show no obvious relation to the function required can be filtered out. The found relations then need to be individually assessed by interpreting the abstract. Information from the abstract needs to be interpreted and compared to the requirements. Subjects can then be further filtered based on the interpretation of the results. If the requirements can not be met the SAO-relation can be removed as a potential means. The remaining means can be further explored by searching an interpreting the full article. The main activity is analyzing the feasibility of the means by extracting more information from the article or other sources if requirements are not answered yet. Eventually a decision needs to be made based on the generated means. We will use a ranked method for valuing the requirement by means of Pairwise Comparison Charts (PCC) [43]. The requirements can be ranked based on the binary scoring of each requirements towards the other. Resulting in a total weight value W_{req} of the requirement. The total value is the sum of the binary test between the *tested requirement* R_t and the *other requirements* R_i and calculated for each individual requirement **(1)**:

$$W_{req} = \sum (R_t > R_i)$$

The means itself will be scored by distributing a fixed set of points over the means per requirement. The means that shows the best performance for that requirement gets the most points, the second the second most points using and so on. We use the following formula **(2)**:

$$S_{MR} = k_M R_i$$

in which S_{MR} is the score of the requirement based on the individual relative value of the mean-requirement score $k_M R_i$. The sum of al means-requirement score is the total dividable points x over the means that are evaluated in that requirement **(3)**:

$$\sum k_{1 \to n} = x$$

The points given per means per requirement are then multiplied with the weight of the requirement calculated based on the PCC in formula **(1)**. The sum of all the scores of the means per requirement comprises the total score. The means that scores best on the total score is the best option. Based on these results further decisions can be made using the following formula **(4):**

$$T_{means} = \sum_{M_1 \to M_n} \left| \sum (R_t > R_i) \right| k_M R_i$$

In which **(1)** is the vector for determining the individual mean-requirement score. The sum of all mean requirement scores make up the total score T_{means}.

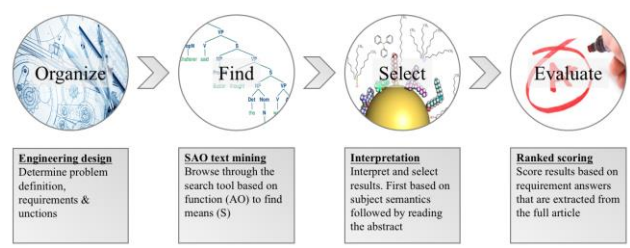


Figure 1 Schematic representation of the methodology (own figure)

IV. Case study

To demonstrate the methodology, we will apply the steps, summarized in *Figure 1*, in an industrial manufacturing firm. The case study must show if the methodology renders results and if these results show to have any value for the firm. It also needs to present if nanotechnology can be browsed as an enabling character that could solve (part of) the innovation problems within the firm. First we will shortly introduce the firm and the case study followed by the execution and the results of the methodology.

A. Introduction of the case study

The firm is a European based product manufacturer and produces large amounts of their products within their own factories in Western Europe. The company needs to innovate to stay competitive and is subject to all kinds of quality standards that are both internal and external. Furthermore, the firm has the challenging aim to reach strict goals in sustainable production and material use. Meaning that many of their materials must be substituted in the near feature if they do not comply with these sustainability rules. Innovation has always been of importance to the company and therefore the organization has innovation highly embedded in their organizational structure. Three main actors are responsible for the organization and execution of the innovation activities. The first actor is the board that sets out innovative goals for the company based on trends, customer and supplier input and technical challenges within the manufacturing and product process. The second actor is the innovation manager that scouts and analyses technologies from outside of the company that could serve the innovative goals. He or she also has the ability to bring new technologies to the board. The third actor is the R&D manager who is in charge of the technical processes. There is a continuous flow of information and feedback between the actors to transfer technical knowledge or to define new innovative goals and tackle technical problems.

The existence of nanotechnology is aware to the company, but the exact implications that nanotechnology could have to their products and processes is not fully investigated. They are interested in further studying the implications of nanotechnology for their current innovative goals to find out if nanotechnology could bring a novel solution. Although they are aware of the availability of science based knowledge and patents the available tools to access this knowledge are insufficient. The information overload and time necessary to understand patens and scientific publications are too large to integrate the scientific scouting of knowledge into the general innovation practices. The company would therefore be greatly helped with a methodology that would help them finding and interpreting scientific knowledge like nanotechnology.

Within the company many targets have been constructed to increase the efficiency and competitiveness of the company. Many of those targets are problems or goals that need to be solved by means of technical innovation. In this matter technical innovation is the optimization or replacement of a component that improves the performance of a process or product. This component may be an innovative solution that can be found in the field of nanotechnology. Or, with other words, a material that has better performance due to its functionality that it gains from nanoscale engineering. To find out if nanotechnology can potentially improve a process or a product within the firm a case has been provided. The case is a technical innovation problem that is, or will be under investigation. The case relates to the flammability control of a textile product. Customer needs and flammability regulations demand a continuous improvement of product fire retardancy while ensuring a better performance. In the following section the case will be more thoroughly explained. Throughout the case the methodology will be applied. This means that first the case will be described according to the problem definition, from this problem definition, requirements will be generated in the form of objectives and constraints. These requirements will shape the solutions space and generate functions. If possible functions will be decomposed to components of the system. Based on these outcomes the SAO search model will be used for generating means. These means will be displayed by a morphological chart and interpreted and evaluated according to the requirements.

B. Problem definition

The company is selling its products all over the world serving a demand in various different industries. An important requirement in many industries is the ability of a product to withstand fire. This ability ensures that the product does not accelerate the growth of a fire and rather slows the spread of the fire down. In order to do so regulations have been designed to control the flammability of products and the formation of smoke when ignited. Products of the firm need to be compliant to these regulations. There are several fire safety levels of which products sold to the aviation and marine industries demand the highest level. Residential and consumer products comply to lower standards.

To be compliant with these levels the firm uses flame retardants in its products. To understand where and how this is embedded in the products a short description of the products will be described. The product consists of multiple layers integrating both textile fibers, and polymeric materials. Flame retardant chemicals are added into the materials to generate a flame retardant product.

Flame retardancy is a point of interest for innovation. This is mainly due to some pitfalls in the use of the flame retardant component which can at a high amount disturb the end-product's function. In addition, reducing cost in flame retardants is a driver. And third, besides regulation concerning heat and fire development there are also regulations concerning the smoke formation during combustion. The current flame retardant, although minimizing heat and flammability, is subject to a lot of smoke during combustion.–Based on the previous we can roughly distinguish three properties the product requires to have:

1. Reduce effect of the retardant material on the end-product performance;
2. Reduce cost of the retardant material;

3. Suppress smoke formation.

For the company there are no strict constraints in the design and application of the fire retardant. If a top coating has a better performance than the current solution they are willing to adjust the production process. They highly value the development of a flame retardant textile product. To achieve the safety standards, they cooperate with a number of external partners.

C. Requirements

The broad problem definition in flame retardancy creates a large solution space. The company is not only looking to phase out the current flame retardant component but also is trying to map the field of fire retardancy with an open minded perspective. How fire retardancy will be embedded in the carpet design is not of importance as long as it fits some general requirements. Because of the open strategy for finding and solving the problem regarding fire retardancy only few constraints have to be generated. Based on the problem definition objectives are designed. Objectives and possible constraints are adapted and listed in **Table 1** These requirements are granted with a measure in order to successfully measure if a scouted technology can be potentially applied as flame retardant.

Table 1 Requirements and measures for fire retardant carpet

Requirement	Type	measure
Can be incorporated in product layer or on product	O	Method of application, when incorporated in fibre no reduction in fibre quality.
Is non toxic	C	No release of toxic materials/pollutants during use and in production
Is allowed as material in C2C list	C	Materials used are not banned
Lowers flammability of the entire product up to the highest standards	C	UL-94 V0
Lowers the smoke formation of the entire carpet up to aero standards	C	*NA*
Is cost efficient	C	Lower cost or is more efficient than current components
Is easy to apply	O	Can be incorporated by the company or company suppliers
Is available and can be tested	O	(commercially) available

D. Functions

The function that the new technology should execute are shown in **Table 2**. Three main functions can be obtained: flammability reduction, smoke reduction and self-extinguishing behavior of the product after combustion. For each of these functions a query is constructed. The query is the input for the search model. The search results will be individually described.

Table 2 Functions and current means of fire retardants

	Function
1.	Reduce flammability of textile product
2.	Reduces smoke formation during combustion
3.	Extinguish fire

Query 1 fire retardant technologies

Action terms (A)	reduce*, decrease*, is
Object terms (O)	flammability, fire, fire-retardant

Query 2 Smoke formation reduction

Action terms (A)	reduce*, decrease*, is
Object terms (O)	smoke, fumes

Query 3 Self extinguishing technologies

Action terms (A)	extinguish
Object terms (O)	fire, flame

For **Query 1** a total of 40 SAO-relation where generated. As the majority of the subjects did not show to be noise every title and abstract was briefly interpreted within the nanotechnology scout. This resulted in an in-depth analysis of 28 SAO-relations. 22 of those SAO relations were after thoroughly reading still a potential means. Eventually 2 SAO-relations could be grouped into 1 technique resulting in 21 potential means. These 21 means were evaluated according to the requirements. **Query 2** did not result in any SAO-relation. This may be the result from a poor query construction, or because there are no sentences with this specific function-means relation. This is imaginable because the structure [SUBJECT-reduces-smoke formation] would render the SAO relation [SUBJECT-reduces-formation] with the algorithm used. **Query 3** resulted in a total of 6 SAO-relations which were all examined based on title and abstract. The 6 results provided two extra means for potential fire retardancy. 5 articles described a layer by layer nanocoating technique and 1 article described a onepot coating technique.

Based on the requirements each technological means will be individually assessed and evaluated. The results are shown in **Table 3**, the information necessary to test the requirements are extracted from the articles that describe the technology.

Table 3 Means requirements evaluation matrix fire retardant carpet

Product / requirements	Method of application	Toxicity	C2C classification	Reduces Flammability	Reduces smoke formation	Availability
Solution 1	Mixed into FR (3,1wt%)	Oxidizing	Not listed	UL 94 V-2	Not reported	Commercially available
Solution 2			Banned			
Solution 3	dispersion of the particles by sonication for 24 h (0,5 – 4 wt%)	Possiby toxic	Not listed	-	Not reported	Available
Solution 4	Synthesis of product and mechanical mixing and heating (>10wt%)	unknown	Not listed	30%-40% heat loss	Not reported	
Solution 5		Possibly toxic				
Solution 6	Melt-blending 5 wt% in several polymers	Non toxic	Not-listed	UL-94 V0	Not reported	Supplied by commercial partners
Solution 7	Melt compounding 10 min in PP (,05-2wt%)	Toxicology assessed	Not listed	20-70%	Not reported	
Solution 8	Dispersed in styrene heavy mixing (3 wt%) swells in styrene	Non toxic	Not listed	Lower heat release rate, faster ignition	Not reported	Commercially available
Solution 9	Melt compounded with a polymer twin screw (4-10 wt%)	Unknown	Not listed	Delayed time to ignition in polymer. 63% reduction in peak HHR	Increased	GCNF was purchased from Applied Sciences Inc
Solution 10	Polymeric blend	After high dose	Material recommended in cosmetics EPEA [44]	Slight decrease in flammability	Not reported	All products were available
Solution 11	Synthesis, mechanical stirring and melt blending (5-10 wt%)	Unknown for Flame retardant	Flame retardant Not listed	27-40% reduction in PHR, up to 20% in THR	Not reported	Raw products are available flame retardant must be synthesized first
Solution 12	Synthesis and mixed screw extruder of LDPE	Low toxicity	Flame retardant not listed	7-51% reduction in in PHR, 20-28% reduction in THR	Not reported	Raw products are available flame retardant must be synthesized first
Solution 13	Intumescent synthesis and mixed in a twin-screw extruder (4,7 to 20wt%)	Unknown for PPSPB	Intumescent not listed	9 to 50% in PHR, up to 25% reduction in THR	Not reported	Raw products are available flame retardant must be synthesized first
Solution 14	1-5 wt% of bentonites in	Not toxic	/	/	/	/
Solution 15	Ultrasonic treatment, stirring and UV curing of samples with 40wt% fire retardant and mineral	Unknown	flame retardant not listed	Up to 57% PHR reduction,	Not reported	Raw products are available flame retardants must be synthesized first

Product / requirements	Method of application	Toxicity	C2C classification	Reduces Flammability	Reduces smoke formation	Availability
Solution 16	Addition of organic nanomaterials (0.5wt%) with flame retardants (5-15wt%) by means of ultrasonic processing and homogenizing	Organic nanomaterials are highly debated in toxicology	Not listed	Up to 67% reduction of PHR	Not reported	All products commercially available
Solution 17	Dip coating layer by layer of flame retardant and mineral was added to the solution.	PAA: Non toxic could irritate skin PEI: Non toxic	Not listed	Up to 72% PHR At 10 monolayers	Not reported	All products are commercially available
Solution 18	Synthesis of dendrimer. In situ polymerization (5wt%)	Unknown	Not listed	Reduction of 58% PHR, intumescent activity	Not reported	Raw products are available, in-house synthesis dendrimer
Solution 19	-	-	Banned	-	-	-
Solution 20	Mechanical stirring of nanomaterials (1hr) and high temperature (4hrs). Suspension polymerization of styrene.	One of the products used is toxic.	Unknown	Peak heat release rate 80% reduction with salts, Flammability class: UL-VI/VO	Less smoke, An specific extinction area reduced with 62%	All products are commercially available.
Solution 21	Layer by layer application of two coating. One is a binder, the second is an intumescent solvent. Requires up to 30 layers	NA	NA	Between 50 and 81% reduction in PHR	NA	NA
Solution 22	NA	NA	NA	81% reduction of peak heat release.	NA	NA

Table 3 shows how the generated means relate to the requirements. It is interesting that based on the scouted technologies the means can now individually assessed based on their requirement performance. The answers for the requirements have all been extracted from the corresponding article. The requirements that couldn't be measured didn't receive any score. Due to the large amount of results it is hard to obtain the best possible options from the table. One way to get more sense of the potential of a means the outcomes need to be scored and compared. This is done by using the PCC and comparative scorecard calculations. The requirements are ranked in the order:

Complies with C2C > Toxicity > Flammability reduction > Smoke reduction > applicability > Price > Availability

The results of the comparative scorecard are shown in **Table 4**. The scorecard shows the ten best results from scoring the technologies. Results with a lot of unknowns automatically become lower in value but retrieving the unknown information can result in a change in the scorecard. Therefore, the scorecard remains a "work in progress" until all the requirements are measured for the technologies or when the users accepts the measurements.

Table 4 Scorecard flame retardancy

Means	Score	Unknowns
Solution 6	157	3
Solution 10	148	1
Solution 20	143	1
Solution 12	133	2
Solution 8	128	3
Solution 17	125	2
Solution 7	116	3
Solution 16	111	2
Solution 1	110	2
Solution 14	106	4

With these outcomes the company could further analyze the results by gathering more information for answering the unmeasured objectives and constraints and then recalculate the score of the results.

V. Validation

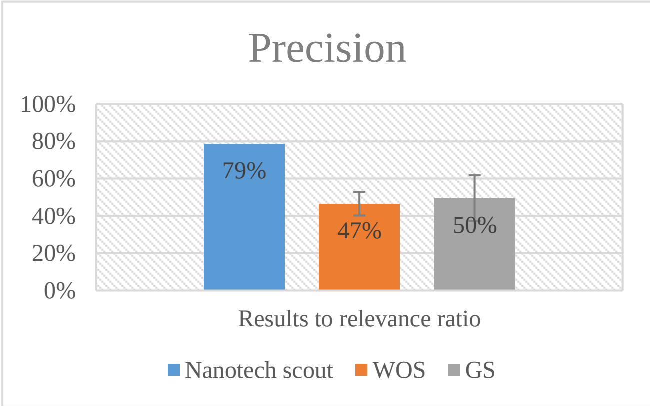
For validating the outcomes, we use the predefined requirements in section *III-A* where the concepts of internal verification and external validation were introduced. For measuring the internal verification of the search model the total amount of SAO-relations in the database are compared to a reference set. Verification showed that 34% of the sentences are converted into SAO-relations. The reference set also showed that the parser for extracting the SAO-relations functions appropriately. For the external validation of the search model the performance of the system is measured using the concepts of precision and recall. Which are often used metrics to measure the effectiveness of information retrieval [40]. The concept of precision illustrates how many of the results are relevant for the case.

$$precision = \frac{|\{relevant\ results\} \cap \{total\ results\}|}{|total\ results|}$$

while the concept of recall describes how many of the retrieved results are part of the total possible results:

$$recall = \frac{|\{returned\ rel.results\} \cap \{total\ rel.results\}|}{|total\ rel.results|}$$

The conventional scientific information retrieval systems Web of Science and Google Scholar were used as a reference. The precision and recall rate of the search model was then compared to the average precision and recall rate of the Web of Science and Google Scholar. The average rate was calculated by constructing 4 queries per retrieval system each and interpreting the first 40 results of the query outcomes. The results are shown in **Figure 2**.



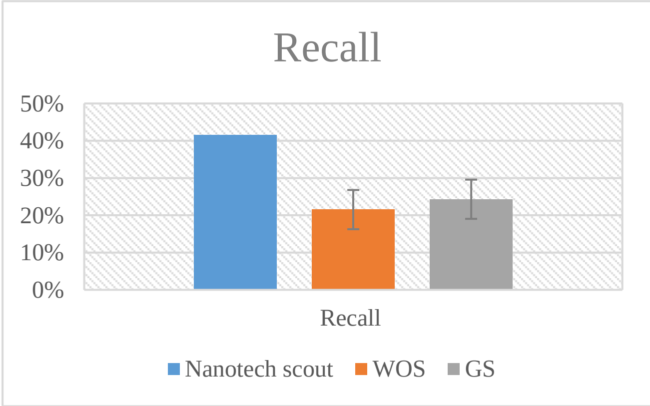


Figure 2 Comparison in precision and recall

The figures show that both in terms of precision and recall the search scout outperforms the conventional information retrieval systems. Validating that both the quality (recall) as the cost (precision) can be lowered by means of the search model as part of the methodology. Within this case the average increase in performance and recall by using the search model as information retrieval system is respectively 63% and 83% when compared to the combined average of Web of Science and Google Scholar.

The search results and the methodology where qualitatively analyzed by means of semi-structured interviews with two R&D managers inside the company. The aim of the interviews was to validate the outcomes, the impact and exploitation of nanotechnology knowledge and the overall viability of the methodology for further use. Both R&D managers confirmed that the results showed new and unknown technologies for solving the innovation problem. The results not only showed completely new technologies, they also showed new insights in solving the innovation problem. With regard to the use of nanotechnology both engineers took a pragmatic stance. For them it does not matter from which science fields technological concepts emerge as long as they can solve the engineering problem they face. Nevertheless, the field of nanotechnology did show new opportunities. Both manager did also value the methodology for future problem solving. One R&D manager valued the time reduction he could realize by applying the methodology and search model. The other manager valued the possibility to find potential solution for a problem individually instead of using external parties to find a solution. This enables better control in the innovation process.

VI. Discussion

The methodology proved to be a useful new approach to scout for science based nanotechnology. The combination of engineering design with SAO-mining is a promising new technique to structure large amount of technical texts. The research also shows that SAO-mining of scientific literature is an interesting new information source for identifying technological concepts besides the use of patents. Because of the flexibility of SAO function-means relation and engineering design the methodology is generalizable to other (bulk) manufacturing firms or generalizable by including/substituting other SAO-indexed databases. Nevertheless, the SAO-mining structure for information retrieval needs to be further investigated. A major challenge is the dependence on the grammatical structure of a sentence. Only 34% of the sentences are converted into an SAO-relation. Meaning that 66% of all the sentences and its information is lost during the parsing process. It can be debated if this loss of information is good or bad, because the sentences do not have an SAO-relation and therefore do not keep a problem-solution relation. However, not all valuable information of text is stored in SAO-relations. For example; the noun flame-retardant also shows flammability reducing behavior. Further analysis is therefore necessary to find out how many valuable information is lost and how this information can be retrieved. Machine learning based on user annotation could be used to find documents and increase the precision and recall rate. Abstracts marked as interesting by the user could be used to mine other abstracts based on similarity.

Another interesting finding is the viewpoint of R&D managers towards the scouting strategy of nanotechnology. Both R&D managers confirmed that they were surprised by the nanotechnology outcomes but both also said that they will not specifically scout for nanotechnologies in a different way. This viewpoint results from the directed and pragmatic search strategies R&D managers have, to solve short term solutions. New technologies should be presented from elsewhere. This is in confirmation with the reasoning by Rafols who argues that nanotechnology suppliers should change their strategy from a material supplier to a solution provider [7]. This would greatly help companies to understand the implication of nanotechnologies for their technological processes.

The complexity of knowledge is also a challenge within the methodology. As described by Hippel information is sticky and adoption of the information only occurs if it is understood by company [45]. Challenges will be in the transfer of the information towards the methodology user with the least amount of information stickiness. Two approaches can be embedded to lower the stickiness of information. The first approach is making the methodology more advanced by automating more steps in the innovation problem solving process. Minimizing the transfer of complex knowledge. The second approach is training the user enabling him or her to apply the methodology in the most optimal way.

Interesting further research can be done in SAO-mining. The technique can be used to explore the nanotechnology field or to explore the performance of SAO in information retrieval and technology scouting in general. The character of nanotechnology can be explored by visualizing what function certain nanotechnologies perform in specific research areas. Building further on the techniques that are already used for exploring technical concepts and [46], [47] technology heat-maps. This information is valuable to improve the understanding on how nanotechnology diffuses into industries, to determine which technological concepts are used where or to measure the most applied technological concepts. SAO-mining itself could be further explored for information retrieval by incorporating patent data and data from various other science fields like chemistry, physics or engineering. The retrieval system could than be further analyzed based on its performance in retrieving information.

VII. Conclusion

In this research a methodology is designed to find, select and evaluate nanotechnology knowledge for innovation problem solving by means of SAO-mining. The methodology is beneficial for scale-intensive production firms to scout and understand nanotechnologies for innovation and problem solving in an efficient manner. Within the methodology the concepts of SAO-mining, engineering design and nanotechnology have been successfully combined. The case study in a scale intensive production firm showed that three carefully designed queries resulted in 22 unique nanotechnology means. The means were interpreted and ranked using information from the abstract and article. External validation confirmed that the performance of the SAO-based information retrieval outranks conventional information retrieval systems. Internal verification showed that almost 2/3 of the information is lost due to SAO-parsing. In future research this loss of information should be further investigated to see if valuable information is lost and how this information could be integrated in the results. Another challenge is information stickiness. We propose two mitigation techniques for solving the stickiness problem; training the user or advancing the automation of the methodology.

VIII. References

[1] T. Nikulainen and C. Palmberg, "Transferring science-based technologies to industry-Does nanotechnology make a difference?," *Technovation*, vol. 30, no. 1, pp. 3–11, 2010.

[2] N. Malanowski and A. Zweck, "Bridging the gap between foresight and market research: Integrating methods to assess the economic potential of nanotechnology," *Technol. Forecast. Soc. Change*, vol. 74, no. 9, pp. 1805–1822, 2007.

[3] M. Meyer, "What do we know about innovation in nanotechnology? Some propositions about an emerging field between hype and path-dependency," *Scientometrics*, vol. 70, no. 3, pp. 779–810, 2007.

[4] S. Wood, R. Jones, and A. Geldart, "The social and economic challenges of nanotechnology," 2003.

[5] J. Youtie, M. Iacopetta, and S. Graham, "Assessing the nature of nanotechnology: Can we uncover an emerging general purpose technology?," *J. Technol. Transf.*, vol. 33, no. 3, pp. 315–329, 2008.

[6] E. Maine, "Scientist-entrepreneurs as the catalysts of nanotechnology commercialization," *Rev. Nanosci. Nanotechnol.*, vol. 2, no. 5, pp. 301–308, 2013.

[7] I. Rafols, P. van Zwanenberg, M. Morgan, P. Nightingale, and A. Smith, "Missing links in nanomaterials governance: Bringing industrial dynamics and downstream policies into view," *J. Technol. Transf.*, vol. 36, no. 6, pp. 624–639, 2011.

[8] G. Scrinis and K. Lyons, "The emerging nano-corporate paradigm: nanotechnology and the transformation of nature, food and agri-food systems," *Int. J. Sociol. Agric. Food*, vol. 15, no. 2, pp. 22–44, 2007.

[9] R. Rohrbeck and H. G. Gemünden, "Corporate foresight: Its three roles in enhancing the innovation capacity of a firm," *Technol. Forecast. Soc. Change*, vol. 78, no. 2, pp. 231–243, Feb. 2011.

[10] E. Lichtenthaler, "Technology intelligence: identification of technological opportunities and threats by firms," *Int. J. Technol. Intell. Plan.*, vol. 2, no. 3, pp. 289–323, 2006.

[11] M. G. Moehrle, R. Isenmann, and R. Phaal, "Technology Roadmapping for Strategy and Innovation," *Charting Route to Success. Berlin al. Springer*, 2013.

[12] E. Lichtenthaler, "The choice of technology intelligence methods in multinationals: towards a contingency approach," *Int. J. Technol. Manag.*, vol. 32, no. 3–4, pp. 388–407, 2005.

[13] P. Atkins and J. De Paula, "Atkins' physical chemistry," in *Chemistry*, Ninth., Oxford, 2009, pp. 783–827.

[14] T. Köhler, S. Mietke, J. Ilgner, and M. Werner, "Nanotechnology – Markets & Trends," pp. 51–56, 2005.

[15] V. Mangematin and S. Walsh, "The future of nanotechnologies," *Technovation*, vol. 32, no. 3–4, pp. 157–160, 2012.

[16] H. Choi and C. C. M. Mody, "The Long History of Molecular Electronics: Microelectronics Origins of Nanotechnology," *Soc. Stud. Sci.*, vol. 39, no. 1, pp. 11–50, 2009.

[17] P. Shapira, J. Youtie, and L. Kay, "National innovation systems and the globalization of nanotechnology innovation," *J. Technol. Transf.*, vol. 36, no. 6, pp. 587–604, 2011.

[18] D. K. R. Robinson, A. Rip, and V. Mangematin, "Technological agglomeration and the emergence of clusters and networks in nanotechnology," *Res. Policy*, vol. 36, no. 6, pp. 871–879, 2007.

[19] A. D. Maynard, R. J. Aitken, T. Butz, V. Colvin, K. Donaldson, G. Oberdorster, M. A. Philbert, J. Ryan, A. Seaton, V. Stone, S. S. Tinkle, L. Tran, N. J. Walker, and D. B. Warheit, "Safe handling of nanotechnology," *Nature*, vol. 444, no. 7117, pp. 267–269, Nov. 2006.

[20] D. B. Warheit, C. M. Sayes, K. L. Reed, and K. A. Swain, "Health effects related to nanoparticle exposures: Environmental, health and safety considerations for assessing hazards and risks," *Pharmacol. Ther.*, vol. 120,

no. 1, pp. 35–42, Oct. 2008.

[21] P. A. Schulte, C. L. Geraci, V. Murashov, E. D. Kuempel, R. D. Zumwalde, V. Castranova, M. D. Hoover, L. Hodson, and K. F. Martinez, "Occupational safety and health criteria for responsible development of nanotechnology," *J. Nanoparticle Res.*, vol. 16, no. 1, pp. 1–17, 2013.

[22] B.-Å. Lundvall, *National systems of innovation: Toward a theory of innovation and interactive learning*, vol. 2. Anthem Press, 2010.

[23] J. Sundbo, *The theory of innovation: enterpreneurs, technology and strategy*. Edward Elgar Publishing, 1998.

[24] C. Freeman and L. Soete, *The economics of industrial innovation*. Psychology Press, 1997.

[25] V. W. Soo, S. Y. Lin, S. Y. Yang, S. N. Lin, and S. L. Cheng, "A cooperative multi-agent platform for invention based on patent document analysis and ontology," *Expert Syst. Appl.*, vol. 31, no. 4, pp. 766–775, 2006.

[26] K. Pavitt, "Sectoral patterns of technical change: towards a taxonomy and a theory," *Res. Policy*, vol. 13, no. 6, pp. 343–373, 1984.

[27] S. J. Kline and N. Rosenberg, "An overview of innovation," *Posit. sum Strateg. Harnessing Technol. Econ. growth*, vol. 14, p. 640, 1986.

[28] H. Chesbrough, "Open Innovation: A New Paradigm for Understanding Industrial Innovation," *Open Innov. Res. a new Paradig.*, pp. 1–12, 2006.

[29] R. Rohrbeck, "Technology Scouting–Harnessing a Network of Experts for Competitive Advantage," *4th Semin. Proj. Innov. 2006; Turku, Finl.*, vol. 40, no. August, pp. 1–15, 2006.

[30] A. Arundel, G. van de Paal, and L. Soete, *PACE report: innovation strategies of Europe's largest industrial firms: results of the PACE survey for information sources, public research, protection of innovations and government programmes: final report*. MERIT, 1995.

[31] A. K. Klevorick, R. C. Levin, R. R. Nelson, and S. G. Winter, "On the sources and significance of interindustry differences in technological opportunities," *Res. Policy*, vol. 24, no. 2, pp. 185–205, 1995.

[32] A. Salter and D. Gann, "Sources of ideas for innovation in engineering design," *Res. Policy*, vol. 32, no. 8, pp. 1309–1324, 2003.

[33] D. N. Mavris and M. R. Kirby, "Technology identification, evaluation, and selection for commercial transport aircraft," 1999.

[34] R. Rohrbeck, "Technology Scouting-a case study on the Deutsche Telekom Laboratories," in *ISPIM-Asia Conference*, 2007.

[35] D. Held, A. McGrew, D. Goldblatt, and J. Perraton, "Global Transformations: Politics, Economics and Culture BT - Politics at the Edge: The PSA Yearbook 1999," C. Pierson and S. Tormey, Eds. London: Palgrave Macmillan UK, 2000, pp. 14–28.

[36] G. G. Chowdhury, "Natural language processing," *Annu. Rev. Inf. Sci. Technol.*, vol. 37, no. 1, pp. 51–89, 2003.

[37] A. Abbas, L. Zhang, and S. U. Khan, "A literature review on the state-of-the-art in patent analysis," *World Patent Information*, vol. 37. pp. 3–13, 2014.

[38] N. Mattas and D. Mehrotra, "Comparing data mining techniques for mining patents," in *Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference on*, 2015, pp. 217–221.

[39] M. G. Moehrle, L. Walter, A. Geritz, and S. Müller, "Patent-based inventor profiles as a basis for human resource decisions in research and development," *R&D Manag.*, vol. 35, no. 5, pp. 513–524, 2005.

[40] C. D. Manning, P. Raghavan, and H. Schutze, *Introduction to Information Retrieval*. New York: Cambridge University Press, 2008.

[41] D. J. Murray-Smith, *Modelling and simulation of integrated systems in engineering: issues of methodology, quality, testing and application*. Elsevier, 2012.

[42] A. P. Sage and J. E. Armstrong, "Introduction to Systems Engineering," *Syst. Eng.*, vol. 13, pp. 3–37, 2000.

[43] C. L. Dym, P. Little, E. J. Orwin, and R. E. Spjut, *Engineering design: a project-based introduction*. Wiley New York, 2004.

[44] EPEA, "50 Recommended Cradle to Cradle Cosmetic Ingredients." .

[45] E. Von Hippel, "'Sticky information' and the locus of problem solving: implications for innovation," *Manage. Sci.*, vol. 40, no. 4, pp. 429–439, 1994.

[46] S. Choi, J. Yoon, K. Kim, J. Y. Lee, and C.-H. Kim, "SAO network analysis of patents for technology trends identification: a case study of polymer electrolyte membrane technology in proton exchange membrane fuel cells," *Scientometrics*, vol. 88, no. 3, pp. 863–883, 2011.

[47] S. Choi, H. Kim, J. Yoon, K. Kim, and J. Y. Lee, "An SAO-based text-mining approach for technology roadmapping using patent information," *R&D Manag.*, vol. 43, no. 1, pp. 52–74, 2013.