Discovering shifts in competitive strategies in probiotics, accelerated with TechMining

Jose M. Vicente-Gomila[1], Anna Palli[2], Begoña de la Calle[2], Miguel A. Artacho[1], Sara Jimenez[2]

[1]Universitat Politècnica de València; Valencia ; Spain

jmvicent@ingenio.upv.es**; +34656804427**

[2]IRTA Institut de Recerca Agroalimentária; Catalunya; Spain

Scientometrics DOI 10.1007/s11192-017-2339-5

Received 11 October 2016

Keywords: Technology Strategy, bibliometrics, animal health, Tech Mining, Semantic TRIZ

MSC 68Q55

JEL C81 or D23

Abstract

Profiling the technological strategy of different competitors is a key element for the companies in a given industry, as well to technology planners and R&D strategists. The analysis of the patent portfolio of a company as well as its evolution in the time line is of interest for technology analysts and decision makers. However, the need for the participation of experts in the field of a company as well as patent specialists, slows down the process. Bibliometrics and text mining techniques contribute to the interpretation of specialists. The present paper tries to offer a step by step procedure to analyze the technology strategy of several companies through the analysis of their portfolio claims, combined with the use of TechMining complemented with a text mining tool. The procedure, complemented with a semantic TRIZ analysis provides key insights in disclosing the technological analysis of some competitors in the field of probiotics for livestock health. The results show interesting shifts in the key probiotic and prebiotic ingredients for which companies claim protection and therefore offers clues about their technology intention in the life sciences industry in a more dynamic, convenient and simple way.

Introduction

Companies try to maintain its profitability and activities by introducing competitive strategies that differentiate them from other competitors either by introducing new products or services to the market or, by reducing costs (Grant 2006; Porter 2008). Appropriability by means of filing patents of a given new technology is one of the main assets for succeeding in innovation (Teece 1986; Kim et al 2016) Through the application of patents, companies can establish a barrier to competitors which results in a competitive asset (Grant 2006; Porter 2008). Nowadays, patents and other intangible assets form a greater portion of the value of a company (Soranzo et al 2016).

By maintaining a strong patent portfolio, life science companies may secure better funding, increase market value or have a driver for gaining or maintaining the competitive advantage of a company (Rose 2007; Abramson 2011). Therefore an increasing interest in analyzing patent portfolios helps companies to anticipate competitors' strategies (Wang et al 2009), and to inform decision makers better. Patent analysis help researchers and policy makers to measure innovation strategies, knowledge diffusion, merger and acquisitions, etc. (Parker et al 2013; Grant et al 2014).

There are many examples showing that patenting key technology achievements represent a big barrier for new entrants. In the case of energy harvesting, some key patents and in particular their claims, could control a whole industry (Kume 2010; Niwa 2016) Energy harvesting is the name for technologies capable of collecting minute amounts of energy from the surrounding. Knowledge intensive industries, as molecular biotechnology or nanotechnology, have also a large patenting activity (Grant et al 2014) In particular, the nutraceutical or medical nutrition industry is also a knowledge intensive industry where patenting is a key element for protecting new products because of a monopoly period of twenty years. (Chih-Hung 2013).

Tracking of patents, patent portfolio analysis and patent landscaping are common strategies to understand the evolution of technologies and the intellectual property strategies of the incumbents, as IP risks and prediction of technology commercialization (Pargaonkar 2016). Patents and intellectual property have moved to the core of the business as we move into a knowledge-based economy (Porter and Cunningham 2005) and into a globalized economy. Most of the patent landscaping work is based on the analysis of fields as title, abstract or citations, and less have concentrated in the claims field. Ferraro et Wanner (2011), and Wang et al (2015) explain about the added difficulty of analyzing claims for a patent landscaping due to the complexity of their linguistic construction. According to Veberne et al. (2010), sentences are longer in claims in relation to normal language, and the syntactic structure and appearance of not common terms make it more difficult to analyze.

There is a need for landscaping patents based on the claims of patents, in a standard and more convenient manner. Traditional text mining patent landscapes which are more focused on trends and foresight try to look for consistent trends in time (Porter et Newman 2011) more than on specific items and their punctual changes in a particular time. Additionally, for the techmining of patent literature purposes, normally due to the extent of full patents, the authors use title and abstracts although recognizing the value of other parts of the patent document (Xie and Miyazaki, 2013; Porter and Cunningham 2005).

The advantage of analyzing claims is twofold. Firstly, it is richer in key terms for the subject of a company and secondly, that terms represent the core elements said company pretends to protect. According to Porter et Cunningham (2005), patent claims contain essential information on the intended purpose of a company. The text in claims is more complete, describing explicitly features in order to provide full legal protection (Noh et al.; 2015).

After the European Union ban of antibiotics as growth promoters (AGP) in animal production, on January 1st 2006 (Animal Task Force 2013), a high interest in alternative methods to improve the performance of livestock has been observed. Furthermore, a significant reduction of antimicrobials as therapeutics has become a principal objective of the EU for the next years. Among the alternative methods there are feed additives with prebiotics and probiotics, phage therapy or bacteriophages, vaccines and mixed probiotics and symbiotic. The latter being a promising alternative (Allen et al 2012).

Probiotics refer to viable microorganisms (bacteria or yeasts) that exhibit a beneficial effect on the health of the host when they are ingested. Prebiotics refer to a non-digestible food ingredient, which are perceived as beneficial to animal health. (Yang *et* Choct, 2009). Symbiotics is a combination of probiotics and prebiotics (Collins et Gibson, 1999; Schrezenmeir et De Vrese, 2001).

The industry of prebiotics and probiotics enter in the feed market and therefore do not follow the long development and test before reaching the market of pharmaceutical companies. This offers a large market opportunity which is being seized with a large patent portfolio (Abramson 2011). However, probiotics are increasingly gaining health applications (Weenen et al 2013), supported by major advances in the identification and characterization of microbes (O'Callaghan et al 2016). Some of these health applications include disbyosis, intestinal diseases of mammals, among others (Foligné et al 2013; Banan-Mwine Daliri et Lee 2015).

The industry of animal health is related to the medical nutrition and therefore it finds itself between the food and the pharmaceutical industry and it represents a fast growing segment within the health and life sciences. As a growing industry, it is relevant to protect their new development through different intellectual property instruments (Weenen et al 2013). Therefore, despite other parts of the patent document could bring a richer description of scope, the protection companies try to accomplish should be clearly stated in the claims of their patents. Visualizing changes in what companies in that industry protect, by means of tracking the claims of their patents, is an object of the present paper.

The present paper proposes a technique for identifying and visualizing changes in the technology strategy of different companies, faster and with less expert involvement by using techmining complemented afterwards with semantic TRIZ. For testing the technique we have applied the procedure in the industry of animal health and more particularly to the industry of prebiotics and probiotics for modulating the microbiome and its application to gut health in livestock and husbandry related animals. The technique of the present paper concentrates on the analysis of the content of patent claims due to their specificity, of at least three competitors in the mentioned sector. The result is then

plotted in the timeline to easily detect if there are strategies which are emerging, or decreasing, if new elements enter into the technology and, if there are disruptions in their patenting intent. It is expected that a shift either in the number or the terms used in the claims may involve either a change in technological interest or a change in resources dedicated to a given technology.

More in detail the procedure performs the following five steps. First, to extract by means of natural language processing the terms of the claims of a searched and selected set of patents; Second, removing common, noisy terms via fuzzy matching, manual touch-up, using term grouping and further applying thesaurus grouping as well as some further clustering scripts; Third, to generate a factors map to extract the main terms in which term clumping strategy also contributed taking into account that newest terms (still nascent) could also be critical; Fourth, plotting such terms in a visualization tool which plots number of citations in claims per year e.g in a Gantt / bubble map , to see the emergence, shifts or disruptions and trends of key terms and classifications; Finally fifth, checking shifts in claim terms being protected and/or comparing the map of different competitors and complementing the explanation of the key observations, with the help of semantic TRIZ answers.

We expect that the proposed claims based patent landscaping complemented with semantic TRIZ, beyond subject-action-object (SAO) based analysis, can be incorporated by companies and technology planers to first quickly detect changes in patent application strategy and to better inform their technical strategy decisions.

The arrangement of this paper performs with the following order. The section Theoretical background presents an overview of scholars' research in the field. The 'Method' section describes the procedure of analyzing claims with the combination of Techmining and Semantic TRIZ. The Case study presents an application of the approach to the emerging field of gut health in livestock as an alternative AGP with the results. The section 'Discussion and Conclusion' discusses the advantages and limitations of the method and outlines future research aspects.

Theoretical background

One of the approaches for studying patent portfolio of a given technology is patent landscaping. Patent maps help to visualize trends and relationships among a group of patents. Patent landscaping can adopt different strategies, including research profiling of patents (Bubela et al. 2013; Grant et al 2014) Patent maps and patent profiling have been object of studies by many researchers and scholars over the last decades (Mogee 1991; Ernst 2001; Chih-Hung 2013; Grant et al 2014; Abbas et al 2014).

Research profiling is one of the five stages in mining external R&D technology (Porter et Newman 2011) Research profiling helps to highlight key outstanding elements. Research profiling uses a modern search engine and text mining tools to characterize large bodies or related literature complemented with the use of TechMining to analyze patent activity of several companies as pattern analyses (Porter et Newman 2011; Porter et Cunningham 2005).

The advent of semantic processing of texts seems to criticize the use of keywords as a valid instrument to extract technology strategies of a company. Although the subject-action-object (SAO) based textual analysis do extract the relationships between words better than just by keywords, such approach is best applied for identifying technological similarities (Park et al 2012; Gerken et Moehrle 2012) competing technologies (Yoon et al 2013) or mergers and acquisitions (Park et al 2013) More specifically, Choi et al (2011), express the limitations of keyword analysis to identify technology trends since keywords cannot express information on how the technology is used or for what purpose. That paper extracts the SAO relationships; however it needs to establish a network to identify key functions based on the number of action-object (AO) relations. The tools using SAO based extraction are limited since they extract also noisy terms (Abbas et al. 2014).

The present paper tries to overcome the limitations of using keywords, by focusing on one critical field of the patent document: the claims. It is in the claims where applicants try to protect their technology strategy. Besides, the present paper focuses only on the earliest patent of the INPADOC patent family, which has claim text in English. It is one of the options of the database used. A patent family is a set of patents filed at different patent authorities that refer to the same invention. It is relevant to select only one family member in order to avoid repeated count of terms in the analysis.

We then apply the techmining technique complemented with semantic TRIZ. To better understand the capabilities of each technique, a short definition of each is following. Tech mining helps to identify trends, technology profiling, and semantic TRIZ allows to understand what the elements of such trends are being used for, what are the causes and effects, components and surrounding elements, etc. The advantage of one aspect of semantic TRIZ is the further interpretation of linguistic forms of different SAO's and the use of other processing tools as mereology and anaphora resolution. It goes beyond the extraction of a list of SAO relationships which need to be statistically treated to offer some meaning. The combination of tools answer to most of the key questions, who, when, where, how, why, etc., an expert may ask him/herself when exploring a technology. The second advantage of semantic TRIZ is its view of technology as an ecosystem and so having into account the relationships of a technology with its surrounding elements. (Vicente-Gomila and Palop 2013; Vicente-Gomila 2014).

Tech mining is a discipline involved in applying information tools to count, interrelate and to analyse science and technology information to help to understand changing and emerging technologies. Tech-mining tools help to analyse, validate and qualifies a large amount of data. They can extract trends, relationships, hidden research networks and weak signals from a vast database of scientific articles. Techmining can use co-words analysis, but it also involves the use of technology measures and innovation indicators to help understand the dynamics of the obtained results and its importance to the subject matter analysed (Porter and Cunningham 2005).

Semantic-TRIZ (Verbitsky 2004) is based on extending the view of systemic functional relationships identified in the TRIZ methodology among the components, features and information related to a technology, with the help of linguistics. Semantic-TRIZ links syntactically and semantically the problem with a solution in a research document or a patent. Implemented in a syntactic-semantic software tool, inspired in the TRIZ methodology, such technology uses linguistic techniques such as mereology, anaphora resolution, cause– effect analysis, applications, failures, properties, etc. to link the functions of a technology, its components and their relationships. It is therefore able to link the knowledge of the user with different – even new– knowledge extracted from documents (Vicente-Gomila and Palop 2013), Techmining uncovers questions as 'what', 'when', where' whereas Semantic TRIZ can add the 'How', 'Why' and 'What for' of a technology being explored. To properly evaluate the strategic technology intent through the patent portfolio, one of the key fields to study is the claims field. It is in the claims where companies define the core invention they pretend to defend of potential competitors (Tong and Frame 1994). Previous research dealing with claims had focused on counting the number of claims being it independent or dependent, may have not been very reliable as a measure of technology relevance (Lanjouw and Schankerman 1999; The UK patent Office patent guide).

The UK patent office patent guide document shows, as way of example, that the same patent (within a family) can have 133 claims when granted by the European Patent Office but only 7 claims when granted by the USPTO or 26 claims when granted by the Chinese Patent Office. Lee et al (2016) use the number of claims, however they propose a stochastic method to have a more accurate inference.

Tong and Frame (1994) demonstrate that the analysis of patent claims may be a better indicator of technological intent or effort. Also, Yang et Soo (2012) coincide on the relevance of the claims in the patent. He et al (2016) also recognize the suitability of using claims and the fewer studies about patents using them. Therefore, in the latest decade, there is a need to research on using claims as better technology indicators. The fact that a different number of claims could distort the number of terms extracted from text analysis, can be overcome by using only one member of a patent family, the earliest of the INPADOC family and distribute the extracted terms of the claims of each unique member across the timeline. The capacity of techmining for identifying trends implemented through a tech mining tool can process the sentences of the claims with natural language processing, extracting the key terms cited in the sentences. In the case presented in this paper, the prebiotics and probiotics as well as coadjuvants for enhancing their effect are the key terms extracted from the claims. Although independent claims can be broad and general however when descending to the dependent claims, applicants need to mention at any level the key probiotic or combination of probiotic and/or prebiotics to be protected, allowing to be extracted and profiled.

Method

The present paper proposes to perform a landscaping based on a research profiling of the patent portfolio of leading companies in the field of livestock health. Said research profiling had to be centred in analyzing the key terms of each of the patent families identified in the portfolio of those companies. To accomplish that, the authors selected to work with the earliest patent member of the INPADOC family of every patent family. Another relevant aspect of the proposed procedure is to focus the analysis in the claims of each patent family. It is in the claims where an applicant discloses the key elements of the technology to be protected. The terms used in claims are key for understanding the strategy in a patent and therefore the technological intent of the applicant.

The research profiling of claims implies analyzing the keywords mentioned in such field of the patent documents either in the first independent claim or in any of the further following claims. Identifying the key terms in the claims and plotting such terms against time allows to understand the trend in key elements, probiotics and prebiotics in the present paper. The monitoring of such strategy for every company in timeline, gives a good perspective for competitors to detect shifts and special movements of the target company.

The advantage of analyzing claims is double. On one hand, it is richer in key terms for a company and second hand, that terms represent the core elements the company pretends to protect. The table 22 shows a sample record of a patent from the portfolio of one of the analyzed companies. As can be seen on table 22, neither the title nor the abstract cite all the terms as in the claims. Even the 'derwent title' which is human composed, cannot be as rich in terms and contents as the claims. By using Techmining and processing through NLP terms, all the terms in the claims, a richer group of terms can be extracted which are core to the strategy of a company. On the table, it can be seen that the Derwent title mentions *N-acetylated oligosaccharide,* and *sialylated oligosaccharide* as key prebiotic terms, whereas the abstract mentions the two terms as well and further mentions a generic probiotic without any specificity.

The claims however, do mention a higher in hierarchy i.e. a hypernym term of the family 'fructo' and 'oligo' for the oligosaccharides mentioned above, and further mention the addition of *arachidonic acid ARA* and *docosahexanoid acid DHA*. Finally, claims further specify the type of probiotic used, *lactobacillus reuteri* or *rhamnosus,* as well as the specific sialylated oligosaccharide: *3'-sialyllactose or 6'-sialyllactose,* and the neutral oligosaccharide *lacto-N-tetraose* in the cited example.

Therefore, it seems clear that the NLP processing of claims can enrich the analysis as compared to title and abstract only analysis and it also can be more specific as more specific terms appear. It is evident that by analyzing the description of the patent, the same type of terms as in claims can appear, however analyzing the full text would include a large quantity of void and jargon type terms. The description part has a more detailed explanaition of the functioning of the technologies involved and about the appearance of the invention (Noh et al. 2015) Nevertheless, the terms appearing in the claims are the core of the protection companies try to pursue in the market space.

Table 22. Extract sample of the claims from one patent of the retrieved set.

Publication number	WO2013057062A120130425 Record 210/1993
Patent Assignee	NESTEC S.A., CH
Title	COMPOSITION FOR USE IN THE PROMOTION OF INTESTINAL ANGIOGENESIS AND OF NUTRIENT ABSORPTION AND OF ENTERAL FEEDING TOLERANCE AND/OR IN THE PREVENTION AND/OR TREATMENT OF INTESTINAL INFLAMMATION AND/OR IN THE RECOVERY AFTER INTESTINAL INJURY AND SURGERY
Title Derwent	Composition used e.g. as supplement, and to treat intestinal inflammation, comprises *long chain polyunsaturated fatty acid*,

	probiotic and mixture of oligosaccharides (having e.g. *N-acetylated oligosaccharide* and *sialylated oligosaccharide*).
Abstract	The invention discloses a composition comprising at least one *long chain polyunsaturated fatty acid*, at least one probiotic and a mixture of oligosaccharides, said mixture containing at least one N-acetylated oligosaccharide, at least one *sialylated oligosaccharide* and at least one *neutral oligosaccharide*, for use in the promotion of intestinal angiogenesis and of nutrient absorption and of enteral feeding toleranceand/or in the *prevention and/or treatment of intestinal inflammation, such as necrotizing enterocolis, and/or in the recovery after intestinal injury and/or surgery.*
Some initial claims	A composition comprising at least one long chain polyunsaturated fatty acid (LC-PUFA), at least one probiotic and a mixture of oligosaccharides, said mixture containing at least one N-acetylated oligosaccharide, at least one sialylated oligosaccharide and at least one neutral oligosaccharide, for use in the promotion of intestinal angiogenesis and of nutrient absorption and of enteral feeding tolerance and/or in the prevention and/or treatment of intestinal inflammation, such as necrotizing enterocolitis, and/or in the recovery after intestinal injury and/or surgery. A composition according to the preceding claim, wherein the neutral oligosaccharide is chosen among *fructooligosaccharides (FOS)* and/or *galactooligosaccharides (GOS)*, preferably GOS. A composition according to any one of the preceding claims, wherein said oligosaccharide mixture contains at least one N-acetylated oligosaccharide selected from the group comprising GalNAcal ,3Gal31 ,4Glc (=3'GalNAc-lac = N-acetyl-galactosaminyl-lactose), Gal31,6GalNAca1,3Gal3"l,4Glc (= 6'Gal- 3GalNAc-lac = galactosyl-N-acetyl-galactosaminyl-lactose),.. A composition according to any one of the preceding claims, wherein the oligosaccharide mixture is present in an amount of 0.5-70%, more preferably 1 - 20%, even more preferably 2-5%, with respect to the total weight of the composition. A composition according to any one of the preceding claims, wherein the LC- PU FA is chosen among *arachidonic acid (ARA)* and *docosahexanoic acid (DHA)*, preferably the *LC-PUFA* is a mixture of ARA and DHA. A composition according to any one of the preceding claims, wherein the probiotic is chosen among probiotic bacterial strains, preferably the probiotic is a lactobacillus o r a bifidobacterium, more *preferably the probiotic is Lactobacillus rhamnosus*, *Bifidobacterium lactis* and *Lactobacillus reuteri*. A composition according to any one of the preceding claims, wherein the N acetylated oligosaccharide is selected from the group comprising lacto-N neotetraose (or LNnT) and *lacto-N-tetraose* (or LNT). 9. A composition according to any one of the preceding claims, wherein the sialylated oligosaccharide is selected from the group comprising *3'-sialyllactose* and *6'-sialyllactose*, and preferably the sialylated oligosaccharide comprises both 3'-sialyllactose and 6'-sialyllactose, the ratio between 3'-sialyllactose and 6'-sialyllactose lying preferably in the range between 5:1 and 1 :2.

The procedure proposed in the present paper takes advantage of the natural language processing capabilities of the techmining tool of Search Technology. The analysis performed by natural language processing, could extract and correlate the key elements of their technology, disclosed in both, the independent claims and the dependent ones. Further analysis of the 1993 patent families with the help of a semantic TRIZ tool, extracted the meaning of any term appearing in the claim, by offering 'applications', 'properties' and definitions' among other typical questions for said term.

In the present paper, the use of claims limits somehow the scope of the semantic extraction since the sentences of the claims are detailed for systemic purpose of affiliation and features, but less for the functioning of the invention. Therefore, to extract the application or uses of the prebiotic and probiotic terms also the description of the patents have been analyzed. This fact does not change the scope of the paper but further reinforces the procedure proposed. The shifts in strategy are analyzed by tech mining, but to further understand the relevance and the context of the terms involved in the shifts of interest, the description of the same set of patents is also semantically analyzed. (Abbas et al. 2014).

The proposed framework consists first in performing a comprehensive search in any existing database using all the keywords and the synonyms and, in the case of animal health, the stemming words of each prebiotic and probiotic class. Grouping the result of the search in patent families and selecting, according to any logic criteria one member of each family to avoid repetition in the number of terms protected by each company. Any patent member in the families, despite of the number of claims, share at least the key elements they want to protect and therefore the key terms mentioning such elements. The next step is to import the results in text mining tool specially inspired in the knowledge of technology management e.g. techmining. By means of cleaning and grouping different variations of probiotic and prebiotic terms, the preparation of the further analysis using different thesauri and using tools as cluster suite script in the techmining tool, techmining tool to remove irrelevant and stop terms.

Next is comparing companies' portfolio by listing applicants and selecting those with a steady activity in the latest 5 years. A key step for advancing strategy analysis is by first grouping elements cited in the claims that act as 'subjects' and grouping elements which may be considered as 'targets' of functions performed by the subjects. Then, a co-occurrence matrix crossing subjects with targets brings an initial clear view of '*who'* is dedicated to '*what'* and what ingredients are targeted to which type of animals as shown in figure 33. The penultimate step is to plot the evolution of different key elements of said claims in time by using, for instance, a bubble chart which shows the number of patents citing any key term distributed by years. It shows the frequency and distribution against time, of the different terms to see which are emerging, which are decreasing or, if there is some disruption in the intent of any company.

Finally, to understand the purpose of different companies in citing key terms in time, the semantic TRIZ tool can bring information about how a term is used, or how is it obtained or what is it part of other components (mereology), or what are the properties exhibited by said term.

The process is summarized in table 23.

Table 23. Steps for detecting shifts in company strategies.

1. A search in in a patent database brought 1993 patent families 2. NLP processing and extraction of claims terms 3. Removing common, noisy terms via fuzzy matching and cluster suite script, both in VP and manual touch up. Using the thesaurus grouping

4. Extracting terms with a factors-map and further selecting with the help of experts from IRTA
5. Establishing a relationship of subjects and objects and cross correlating them with the help of the co-occurrence matrix
6. Going further the comparison of terms but also adding the dynamics of such terms against its distribution in time to see the dynamics of key elements in the claims and, to see the shifts and changes in time
7. Take advantage of semantic processing to add meaning about the uses and properties of any of the components in the claims

Figure 33 shows an example of step 5 in the table, using a co-occurrence matrix crossing the subjects, e.g. probiotics, prebiotics, or special food ingredients as fish oil, in the present paper and targets, e.g. animals or a specific part or aspect of an animal or a pathogen present in livestock animals. Such co-occurrence establishes a SAO relationship based on techmining with statistically relevance, therefore outperforming some of the limitations of the traditional SAO analysis.

Reset		Abstract (NLP) (Phrases) (Cleaned): e	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
		# Records	92	68	65	63	58	56	55	47	44	43	43	41	33	31	30	29	27	19	17	17	15	14	14	14
d): except stopwords (Cleaned) (Cleane	# Records	Show Values >= 1 and <= 50 Cooccurrence # of Records	veterinary composition	antibody	fish oil	reducing methanogenesis , feeding and an	probiotic bacteria or compositions	antibiotic	polypeptide	vitamin	antigen	microorganism	peptide	lactic acid bacteria	diet	amino acid	fatty acid	Bacillus	dietary supplement	polynucleotide	carotenoid	zinc	bacteriophage	astaxanthin	Bacillus subtilis	lactobacillus
1	41	fish		10	3		10	6	6	6	9	9	6	2	5	4	4	5	5	1	9		1	7	4	1
2	31	poultry	1	3	1	2	8	14	4	4	3	8		9	7	2	1	6	4	3	5		8		8	3
3	31	livestock	1	1		4	8	12	1	5	4	3	5	7		2	1	1			2	2	5		1	
4	26	pig & piglet & swine	2	6			6	6	5	4	4	1	2	2	4	2		2		4		1				2
5	25	ruminant	1	3	2	50	2	2	1	6	1	5	2	4	9	5	6	1	1			1	1		1	
6	19	cattle or cow	2	5	2	1	4	1	3	7	2	5	1	3	1	2			1	1		2		1		1
7	16	gastrointestinal tract		1		7	9	4	1	4		4	1	7	6	2	3	3	3		1	1	4		2	5
8	11	mammal	3	5	1	1	3	1	1	2	8	2	4		1		1	1	2	2	1			1		1
9	72	bird	4	1			1	3	1	1	2		1	3	2		1	2	4							2
10	60	nematode	2	2							1															
11	54	sheep		3		1	1			5	1	1			1	1				1						
12	52	improving ruminant animal performance				15									2		1									
13	51	egg	1	3	1			3			3			2	3				2		1		1			
14	39	domestic animals	1	1				2								1		1	2							1
15	38	inflammation	1		1	1		1											1							
16	35	bovin	1	1			1	1	1	2		1	1													
17	34	goat		3	1	1				1		1							1	1						
18	33	crustacean		3			7	1		3	2	1			1		1	7			1		1			
19	31	pathogen		2			1	1			2					1		1								1
20	31	virus		1					1		2									1						

Figure 33. Example of co-occurrence matrix showing the ingredients as subjects and the animals or aspect of animal's f key elements in patent family claims.

Results

The method outlined before has been applied to analyze the strategy of several competitors in the market of prebiotics and probiotics for the gut health of livestock. In this industry where the market is

expanding, the protection of new ingredients or the new combination of ingredients open opportunities to step into said market.

To test the combination of the two techniques, a techmining exercise was developed following the nine steps of the decision phases and the techmining process (Porter and Cunningham 2005).

A search in a patent database with output of data from one member of each patent family selected according to the criteria previously described, brought 1993 patent families, including information about years, inventors, assignees, title, abstract and claims. With the natural language processing of a techmining tool, the claims were processed and a group of more than 150,000 claims NLP phrases were extracted.

From them, using thesaurus grouping, expert help and cluster suite script for selection of terms and removal of void terms, reduced the number of valid terms from which the probiotic and prebiotic terms were grouped and classified.

The authors then analyzed the assignees to select the most dynamic applying for patents in the later years and the key terms in the claims of said patents, in order to understand how the animal medical nutrition industry is focusing its activity. The search in Thomson Innovation brought about 1993 patent families with a time span of year 2000 through 2015, who are indeed dedicated to animal gut health and not mainly to humans. From which, as a way of example, three of the competitors or players selected either by being active in the later years or by being relevant to the experts, as Nestec, the 'Institute de Recherche Agroalimentaire' INRA and DSM, have been analysed according to the steps described before. The results show that by means of visualizing the trend and intensity of the patent claim terms protected in time and specially any shift in said terms being protected, means a shift in the interest of any player. Further understanding what and how said terms are being used in a faster way, combining techmining and semantic TRIZ, can be very convenient to decision makers.

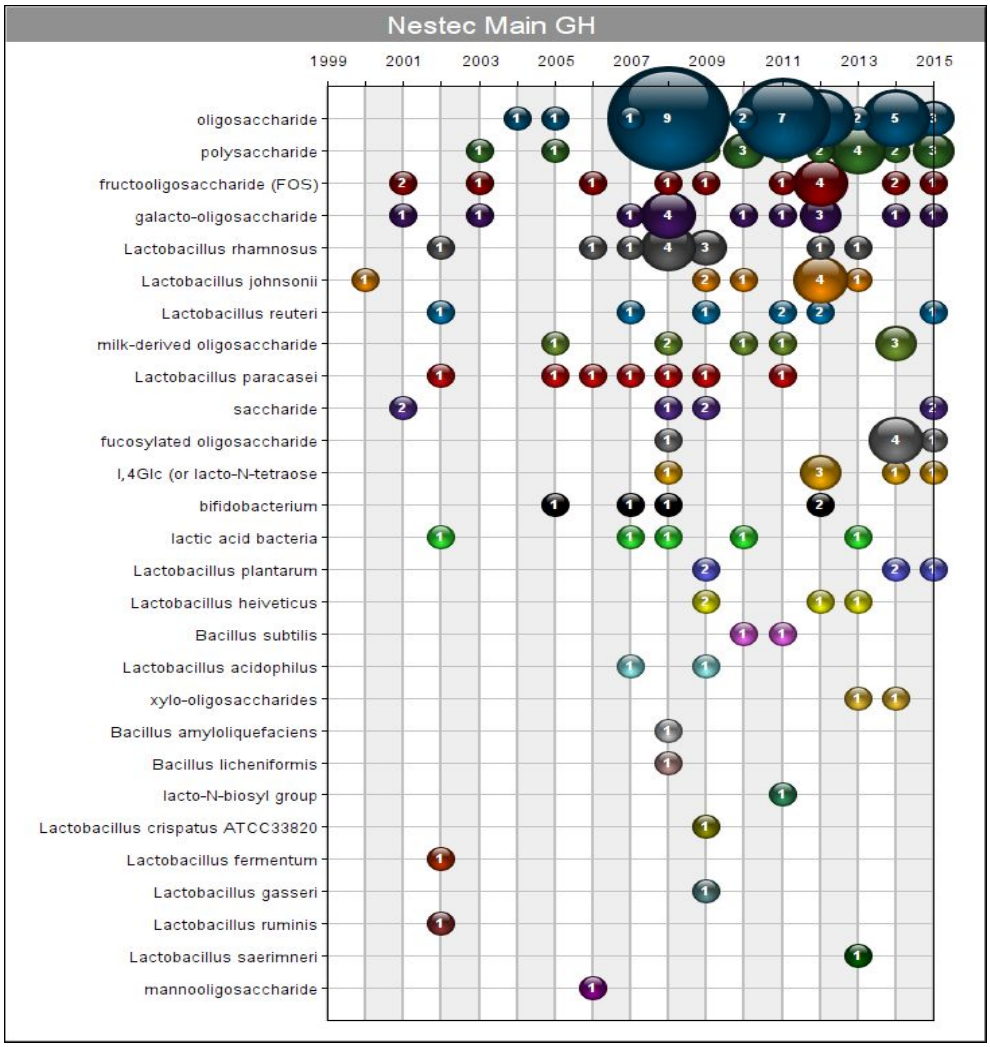

Figure 34. Bubble map with trends about health agent terms in claims of Nestec.

Figure 34 shows a main activity of the company Nestec and its remarkable concentration on the prebiotic components of the family of 'oligosaccharides'. Prebiotics are key since they modulate or change the ratio between 'good' bacteria and microflora and 'bad' bacteria. Prebiotics are mainly used as 'food' for the gut microflora and so its specificity favours some bacteria instead of others. Returning to the figure 34, a decreasing effort in nonspecific oligosaccharides and a surge of interest in one particular 'fucosylated oligosaccharide' is shown. The authors have seen many of the patents trying to protect the combination of some probiotics with some prebiotics.

'Polysaccharides' and 'lactobacillus johnsonii' seem also to catch some attention. Tables 24 and 25 show some applications of lactobacillus johnsonii and fucosylated oligosaccharide, extracted with a syntactic semantic tool. These applications offer some ideas about how both are being used and therefore offer a guide about their significance. Said applications therefore, offer insights about what could be a probable reason for Nestec to reduce in general oligosaccharide and, to start or to concentrate in fucosylated oligosaccharide. From the use or applications, complemented by expert advice, it seems that this company is positioning in preparing the animal gut. It could point to a strategy which is to focus in the surroundings which may host the probiotics, protected by itself or by other incumbents.

As these terms are extracted from the claims, the presence of such ingredients means the concretion about what the company may be protecting. The next point would be the exact amounts, combinations etc. of ingredients which are beyond the scope of the present paper, but which are also key to the protection of said company.

Table 24. Some key applications of *lactobacillus johnsonii* as a health agent extracted from the set of patents.

Lactobacillus johnsonii uses or applications	
Inhibition of intestinal invasive pathogen	(5) instances
Inhibition of enterotoxin	(5)

Table 25. Some key applications of *fucosylated oligosaccharide* as a health agent extracted from the set of patents.

Fucosylated oligosaccharide uses or applications	
As a competitor for pathogens	(2)
Inhibition of stable toxin of Escherichia coli	(2)
As prevention of influenza infection	(1)
For combating multiresistant bacteria	(1)
As health promoter	(1)

The next case is about an Institution present in the market of gut health and more concretely in probiotics and prebiotics for the health of mammals, including humans, as many companies try to reach as broad as possible the potential applications in the medical nutrition industry. This institution was of interest to the experts interested in the present study.

The figure 35 shows some interesting shifts about the claims of the *Institute National de la Rechereche Agroalimentaire*, INRA. In 2003 and 2012-2013 INRA has protected about polysaccharides in the prebiotics. In 2009 and 2010 INRA protected some use about '*lactobacillus sakei*' but in 2014 and 2015 they protect many lactobacillus strains except the lactobacillus sakei. This could mean either they see more interest in other lactobacillus or they have enough protection for lactobacillus sakei. The bubble map does not show clearly any strategy; however, it may call the attention of a subtle change or disruption in a graphical and notorious way. The chart shows only one citation per year due to a comparatively smaller number of patents than the other selected companies. The chart, however, shows a clear shift from protecting the prebiotics 'oligosaccharide' and 'polysaccharides' during the period of 202 to 2014, to a number of different probiotics in 2015 and none prebiotics. By number it may not show a clear trend but in comparing their own portfolio can show a shift in their strategy.

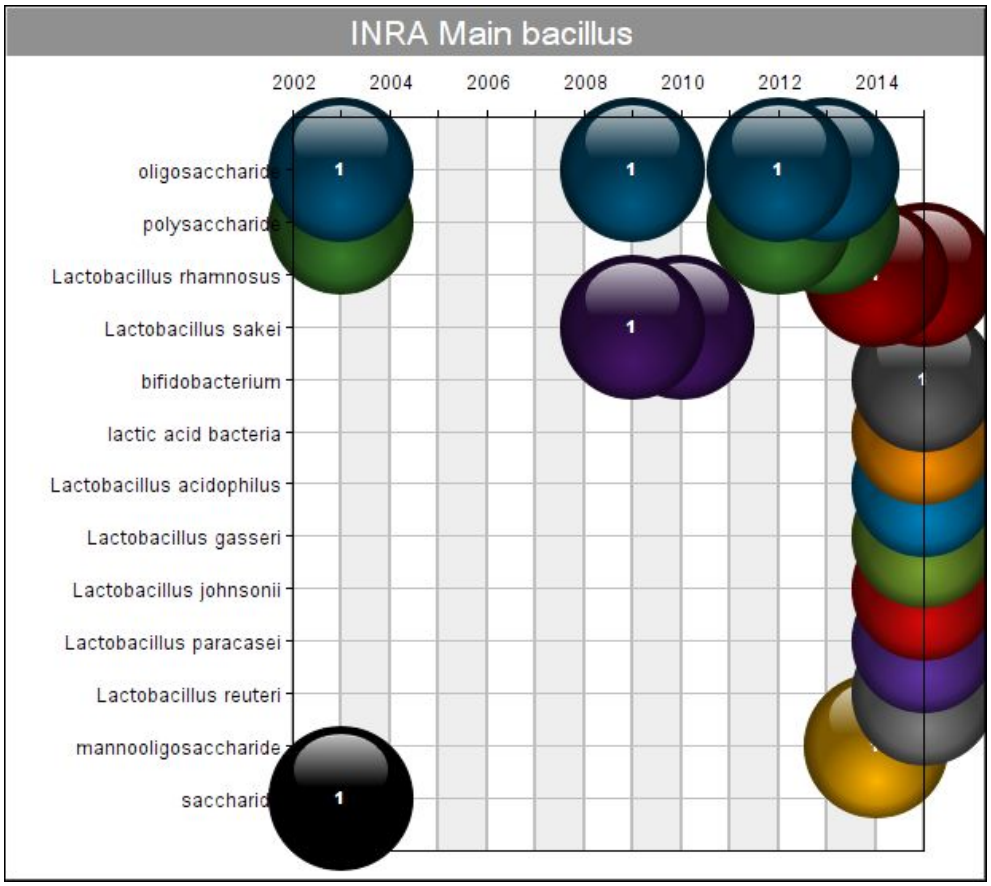

Figure 35. bubble map with trends about health agent terms in claims of INRA.

Table 26 shows some properties about *lactobacillus sakei*. It is worth to mention the high *phytase* activity which is an interesting trend since it may help to reduce the burden of contamination of livestock industry by reducing the amount of phosphorus residues (Kaushik 2015)

Also to mention in the bubble chart about *lactobacillus rhamnosus*, whose applications are shown in Table 27. To mention the effect against yeast and Escherichia coli, two worrying items for the livestock gut health.

Table 26. Some key applications of lactobacillus sakeii as a health agent extracted from the set of patents.

Lactobacillus Sakei uses or applications	
High phytase activity	(4)
Best antibacterial property	(3)
Improves infection resistance for E. coli	(2)
Produces sakacin A effective against listeria	(2)

Table 27. Some key applications of lactobacillus Rhamnosus as a health agent extracted from the set of patents.

Lactobacillus Rhamnosus uses or applications	
Against Escherichia coli	(5)
Against streptococcus mutant	(5)
Against streptococcus sobrinus	(5)
Inhibition of pro-inflammatory cytokines	(5)
Cure of yeast infection	(4)

Figure 36 shows the bubble chart of DSM IP and its activity in protecting probiotics and prebiotics, mainly in the latter. To mention a great surge in 2012 and 2013 in the prebiotic polysaccharide and a stop in 2010 about oligosaccharides and galacto-oligosaccharides. The applications of the polysaccharide are shown in table 28. Dietary fibre is among the expected uses of any fibre, however the antiviral and antitumor may be less common properties. As in the case of Nestec, this group of big players in the animal health industry seem to concentrate in the surroundings, in this case, the preparation of the animal gut lining, maybe to be able to protect later possible probiotics in the future.

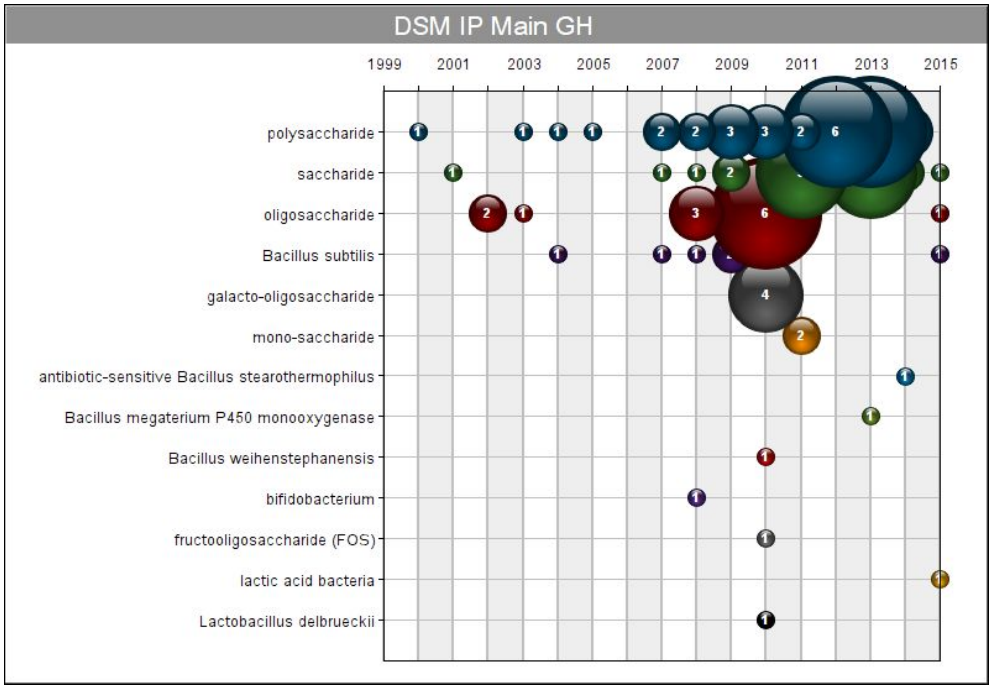

Figure 36. bubble map with trends about health agent terms in claims of DSM IP.

Table 28. Some key applications of the prebiotic family of polysaccharides as health agent extracted from the set of patents.

Polysaccharide applications in animal gut health	
For antibody generation	(20)
As antitumor agent	(12)
As antiviral agent	(11)
As pneumococcal vaccine	(11)
As dietary fiber	(10)

Discussion and Conclusions

The framework outlined here brought in a visual form and using standard procedures, a method to quickly detect changes or small disruptions in the protection strategy of competitor firms. The combination of a techmining approach, complemented further with that of semantic TRIZ, can propose a simple, yet direct, method to detect changes in the strategic intent of different companies, reflected in shifts or continued effort through the terms extracted from their claims in their patent portfolio and quickly seeing the purpose of such terms through their uses or applications extracted from the claims and complemented with those of the descriptions.

Having analyzed the claims of the patents in the nutraceutical brings advantages in front of traditional title/abstract approaches. Finding the key terms in the claims and their pattern against time, is a direct indicator of what companies are trying to protect and therefore central to the strategy of said company. Furthermore, as we have outlined that claims are richer in the key terms and very precise in citing such key terms with respect to titles and abstracts, the proposed method brings closer to detect any new ingredient or coadjuvant a company tries to protect. In conclusion, we may say that by mining claims we analyzed more and more relevant terms related to the strategy of a company.

The help of the semantic analysis quickly helps to understand the context and the purpose of the new terms signaling the direction of the patent strategy of a company, with less expert involvement. The different number of citations in the uses or applications is due to the number of patents analyzed for each company and to the extent of prebiotics as polysaccharides which are combined with the claim of new probiotics.

In the present study, each company had their own participation in the nutraceutical industry and so the number of patents each company has filed. The scope has been to analyze each company isolated to detect changes in its patenting behavior. An interesting and complementing approach could be to compare dynamically the strategy of all participants in the industry to detect which is expanding into the white space of others or vice versa.

The present framework needs further investigation. For instance, it is worth to check the application of the present method to different technologies of other industries different than in the medical nutrition industry where expanding the protection of substances as an alternative to antibiotics is the main strategy. In this industry, the market is expanding, therefore the protection of new ingredients or the new combination of ingredients, open opportunities to step into said market.

For instance, in the hardware industry, it is necessary to check if the shift of claim terms in time, corresponds to relevant positions or a change in the strategies of involve companies. Maybe in mechanical or electronic domains where new components may have more or less impact in the opportunities of the company, further dimensions should be considered. Another element to further explore, is the use of other visualization tools using also information about the market to better understand the possible movements in the protection through patents.

A more comprehensive study could be performed comparing the extraction of terms from claims compared to other parts of the same patent portfolios of the selected companies. One limitation of the present study is the lack of a longer period or to be able to explore if the current changes in strategy are steady in the future. Another limitation is the scarce involvement in patenting of some companies in the analyzed domain. Companies are trying to bet their presence in the market space, by protecting now what is thought to be of real value in the near future.

The present study helps to detect changes in the strategy, further analysis should be carried to fully understand the impact of any change in the strategy of the involved companies. The study about priorities and countries covered should also be added for analyzing and understanding any change in the technology strategy.

Aknowledgements

The authors would like to thank the contribution of the research institute IRTA, to the TRIZ company triz XXI and to Fernando Palop and their wise insights and guidance

The authors thank the usage of Search Technology's VantagePoint and IHS-Markit's Goldfire

References

Abbas A.; Zhang L.; Khan S.; (2014). A literature review on the state-of-the-art in patent analysis. World Patent Information; 37; 3-13

Allen, H.; Levine, T.; Bandrick, M; Casey, T.; (2012) Treatment, promotion, commotion: antibiotic alternatives in food-producing animals; Trends in Microbiology; vol. 21; nº 3

Animal Task Force (2013). Research & innovation for a sustainable livestock sector in Europe, [http://www.animaltaskforce.eu/Portals/0/ATF/horizon2020/ATF%20white%20paper%20Research%20priorities%20for%20a%20sustainable%20livestock%20sector%20in%20Europe.pdf] accessed day 4/september/2016

Abramson, D. (2011). Patent strategies for life sciences companies to navigate the changing patent landscape; Journal of Commercial Biotechnology; 17; 358-364

Bubela, T.; Gold, R.; Gregory, G.; Cahoy, D.; Castle, D. (2013). Patent landscaping for life sciences innovation: toward consistent and transparent practices; Nature Biotechnology 31, 202–206

Chih-Hung H. (2013). Patent value assessment and commercialization strategy; Technology forecasting & Social Change; 80; 307-319

Collins, M D.; Gibson G. (1999). Probiotics, prebiotics, and synbiotics: approaches for modulating the microbial ecology of the gut; Am J Clin Nutr 1999;69 (suppl):1052S–7S.

Ernst, H. (1998). Patent portfolio for strategic technology management; Journal of Engineering Technology Management 15; 279–308.

Ferraro, G.; Wanner, L. (2011). Towards the derivation of verbal content relations from patent claims using deep syntactic structures; Knowledge-Based Systems 24; 1233–1244

Foligné, B.; Daniel, C. ; Pot, B. (2013). Current Opinion in Microbiology 2013, 16:284–292

Gerken, J.; Moehrle, M. (2012). A new instrument for technology monitoring: novelty in patents measured by semantic patent analysis; Scientometrics 91:645–670

Grant, E.; Van den Hof, M.; Gold, R. (2014). Patent landscape analysis: A methodology in need of harmonized standards; World patent Information; 39; 3-10

Grant, R. (2012). Contemporary strategic analysis; 5th edition; ISBN 1-405-1999-3

IHS Goldfire; www.ihsmarkit.com (accessed November 2016)

Kim, B.; Miller, D.; Mahoney, J. (2016). The impact of the timing of patents on innovation performance; Research Policy 45 (2016) 914–928

Kaushik, G. editor, (2015). Applied environmental biotechnology: present scenario and future trends; ISBN 978-81-322-2122-7; Springer Verlag

Kume, H. (2010). From low power to no power through energy harvesting: powering up the battery-free world; Nikkei Elctronics Asia; October 31 2010; accessed November 2011

Lanjouw, J. and Schankerman, M. (1999). *The quality of ideas: measuring innovation with multiple indicators.* 7345. National Bureau for Economic Research, Cambridge, MA., USA.; http://www.nber.org; accessed Sept 2016

Lee, C.; Kim, J; Kwon, O.; Woo, H.G. (2016). Stochastic technology life cycle analysis using multiple patent indicators; Technological Forecasting & Social Change 106 (2016) 53–64

Niwa, S. (2016). Patent claims and economic growth; Economic modelling; 54; 377-381

Noh, H.; Jo, Y; Lee, S. (2015). Keyword selection and processing strategy for applying text mining to patent analysis; World Patent Information; 42; 4348-4360

Pargaonkar, Y. (2016), Leveraging patent landscape analysis and IP competitive intelligence; World Patent Information 45; 10-20

Park, H.; Yoon, J.; Kim, K. (2012). Identifying patent infringement using SAO based semantic technological similarities; Scientometrics 90; 515-529

Park, H.; Yoon, J.; Kim, K (2013). Identification and evaluation of corporations for merger and acquisition strategies using patent information and text mining; Scientometrics 97:883–909

Porter, M. (2008). The five competitive forces that shape strategy; Harvard business Review; January 2008

Porter, A. et Newman, N. (2011). Mining External R&D; Technovation 31 ; 171-176

Porter A. L., Cunningham, S. (2005). Tech Mining; Wiley Interscience

Regulation (EC) No 1831/2003 of the European Parliament and of the Council of 22 September 2003 on additives for use in animal nutrition Regulation (EC) No 1831/2003 of the European Parliament and of the Council http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32003R1831

Schrezenmeir J.; De Vrese, M. (2001). Probiotics, prebiotics, and synbiotics—approaching a definition; *Am J Clin Nut* February 2001 vol. 73 no. 2 361s-364s

Soranzo, B.; Nosella, A.; Filippini, R. (2016). Managing firm patents: A bibliometric investigation into the state of the art; J. Eng. Technology Management; Article in press

The patent guide; a handbook for analyzing and interpreting patent data. UK Intellectual patent office

VantagePoint, www.theVantagePoint.com (accessed 20 September 2016).

Verberne, S.; D'hondt, E.; Oostdijk, N. (2010). Quantifying the challenges in parsing patent claims, in: the 1st International Workshop on Advances in Patent Information Retrieval (AsPIRe'10), March 2010, Milton Keynes, UK.

Vicente-Gomila, J.M.; Palop, F. (2013). Combining Combining tech-mining and semantic-TRIZ for a faster and better technology analysis: a case in energy storage systems; Technology Analysis & Strategic Management; Vol. 25, No. 6, 725–743

Vicente-Gomila, J.M.; (2014). The contribution of syntactic-semantic approach to the search for complimentary literatures for scientific or technical discovery; Scientometrics; DOI 10.1007/s11192-014-1299-2

Wang, M.; Chiu, T.; Chen, W. (2009). Exploring Potential R&D Collaborators Based on Patent Portfolio Analysis: The Case of Biosensors; PICMET 2009 Proceedings, August 2-6, Portland, Oregon USA

Wang, J.; Lu, F.; Loh, H. (2015). A two-level parser for patent claim parsing; Advanced Engineering Informatics 29; 431–439

Weenen, T.; Pronker, E.S.; Commandeur, H.R.; Claasen, E. (2013). Pharmanutrition 1 13-21

Xie, Z., & Miyazaki, K. (2013). Evaluating the effectiveness of keyword search strategy for patent identification. World Patent Information, 35(1), 20–30.

Yang, P.A.; Choct P. (2009). Dietary modulation of gut microflora in broiler chickens: a review of the role of six kinds of alternatives to in-feed; Antibiotics; World's Poultry Science Journal, Vol. 65, March 2009

Yoon, J. Park, H.; Kim, K. (2013). Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis; Scientometrics; 94; 313-331