Leverage Patent Analytics to Achieve Business-Oriented Objectives: A Pragmatic Approach

B. Chevalier
Questel,
Grenoble, France, bchevalier@questel.com

Introduction. Because patent publications are at the forefront of emerging technologies and are related to technologies with commercial potential, many companies consume patent landscapes or analytics to get more information, more data on competitors and cleverly construct incremental improvement. Nevertheless, it is not enough to anticipate technological changes and a true structuration of the information is essential to identify business markers and trends. Methods. We have used a big data and data mining approach to process patent information and determine weak signals and market shifts. The following process has been followed: mapping, categorization according to a taxonomy, business markers and trends identification. The domain of AI in medical devices has been studied to illustrate the method. Maps are used for simplifying the data analysis by leveraging keywords identified by semantic algorithm. Considering the volume of the topics (macro, meso, micro) the analysis will be adapted to get certain insights. To pass from maps to categorization analysis we have set up a taxonomy, based on the knowledge of experts and previous data mining work, which allows us to search for non-obvious solutions and objectively focus our attention on all the segments. Supervised machine learning methods help to distribute documents according to taxonomies. Then, maturity and aggressiveness can be qualified based on IP events such as litigations, licensing actions, growth rate or number of applicants. The last step is related to the essence of a landscape, interpreting any weak signals to anticipate the future success. Results and Discussion. We have focused on recent patents, deserted areas on the map and the taxonomy and on analyzing “unusual” patent proceedings to determine new R&D directions and innovation pathways for the use of AI in medical devices. Conclusion. We have found it particularly relevant to use taxonomies and IP events landscape of patents to anticipate technological trends and market directions and we are convinced that the sophistication of AI-based solutions will push the predictions of the markets further.
Введение. Инновационные технологии, имеющие коммерческий потенциал, напрямую связаны с патентными публикациями, в связи с чем многие компании используют патентные ландшафты или патентную аналитику для получения более подробной информации о рынке и конкурентах, а также проработки стратегии развития. Однако для реализации стратегических целей компаний недостаточно прогнозирования технологических изменений – необходимо правильно структурировать информацию для выявления бизнес-маркеров и тенденций.

Методы исследования. Для обработки патентной информации, определения слабых сигналов и структурных изменений рынка использованы методы больших данных и интеллектуального анализа данных. Построены патентные ландшафты, затем проведена категоризация в соответствии с таксономией и выявлены маркеры и основные тенденции развития бизнеса.

Результаты и дискуссия. В качестве иллюстрации эффективности выбранных инструментов и стратегии была изучена область применения искусственного интеллекта в медицинских устройствах. Патентные ландшафты использованы для упрощения анализа данных с помощью ключевых слов, выявленных семантическим алгоритмом. Результаты анализа адаптированы с учетом объема выборки (макро-, мезо- и микроуровни). Методы машинного обучения использованы для рассмотрения патентных документов в соответствии с таксономией, позволяющей находить нестандартные решения и объективно оценивать все сегменты данных. Степень проработки (зрелости) и конкурентоспособности технологии определены на основе анализа событий на рынке интеллектуальной собственности, таких как судебные процессы, предоставление лицензий, динамика темпа роста или количества заявителей. Последний этап исследования посвящен анализу патентного ландшафта посредством интерпретации сигна-
лотов рынка для выявления перспективных направлений. В исследовании особое внимание уделено патентам последних лет, неисследованным областям патентной карты и таксономии, а также аналитике «необычных» патентных процедур для определения новых направлений научно-исследовательских и опытно-конструкторских работ и инновационных путей использования искусственного интеллекта в медицинских устройствах. Заключение. Сделан вывод о важности использования таксономии и патентного ландшафта для прогнозирования технологических тенденций и направлений развития рынка. Повышение сложности решаемых задач с учетом использования искусственного интеллекта будет способствовать дальнейшему развитию методов прогнозирования.

Ключевые слова: патентная аналитика, наукометрия, интеллектуальная собственность, патентная информатика, патентный ландшафт, построение патентных ландшафтов, тенденции развития технологий, искусственный интеллект, кластеризация, интеллектуальный анализ данных

Для цитирования: Шевалье Б. Использование патентной аналитики для достижения бизнес-ориентированных целей: прагматический подход // Управление наукой и наукометрия. 2020. Т. 15, № 2. С. 120–135. DOI: https://doi.org/10.33873/2686-6706.2020.15-2.120-135

Introduction

As an entrepreneur, R&D manager, CTO, or Chief IP Officer, sooner or later you will eventually face a shift in your sector. This could be a slow and latent move, or a disruptive and brutal change, but it will force you to make the right IP and innovation decisions to keep your company moving in the right direction.

This is hard to do without a clear overview of your industry; who is doing what and since when. Therefore, patent analytics and IP landscapes have become so popular and common. Because patent publications are at forefront of emerging technologies and are related to technologies with commercial potential, many companies consume these reports to get more information, more data on competitors and cleverly construct incremental improvements. Nevertheless, is it enough to anticipate these technological changes? Is patent aggregation alone enough to decode the weak signals sent by the market? Can you decide between focusing on your own R&D or acquiring external technologies?

With more than 120 million of patents published, digitized and structured in huge patent database software, the question is not, do we have the data? But what can I do with it and how?

We will lead you to an answer by first introducing how crucial the mapping phase is to exploit the data and make accurate assumptions and in-
sights. A landscape must provide clear indications of the maturity and the trends of your market to eventually tell you where you need to go, which innovation pathway to take, and what common mistakes to avoid.

After which, as the architect Cedric Price said, “Technology is the answer, but what was the question?” The final conclusions of a landscape must carve the way of the technology appropriation. From the results, you will know if you are going to scout for a company, acquire licenses, launch a freedom to operate study or accelerate the pace of your R&D team project.

**Literature Review**

Patent analytics describes the science of analysing large amounts of intellectual property information, in relation to other data sources, to discover relationships and trends [1].

Literature converges on the fact that patent is high-quality data that can be used for decision making on a strategic level in all kind of organizations [2]. Patent visualization and road-mapping are essential to define topics of interest and possibility to protect rights [3]. But patent data analytics is also a field where new information technologies, like artificial intelligence and blockchain, can easily be adopted and create potential breakthrough progress [4].

**Methods**

Patent analytics differs from traditional statistical approach of analysis. Due to the huge number published documents, it is close to big data analytics and data mining solutions [5–6]. It all starts with data that you need to mine, clean and model in a way to explain a response. Hence, the process flow looks like what is presented in Fig. 1.

**Results and Discussion**

A landscape is the process of collecting documents and building a data-set for different analytics and metrics to be computed. The geographical coverage is necessary to explore the R&D locations and be aware of where the IP rights exist, and the dynamics of the domain, players, patent classifications and concepts are mandatory. As a rule, a landscape report is also enriched with details about which documents are the most important (in the case of patents, we will focus on the highest cited documents), where assignees have worked together (co-assignments, co-publications), what are the actual statuses of the documents (alive? published?) and which inventors are the most prolific.
Mapping

In its purest definition a map is a visual representation of a domain, process, structure, or a system that depicts the arrangement of and relationships among its different components. By understanding how a component influences or is linked to another, strategies are unveiled, critical components are pinpointed and flows of information traced [7].

With 2.9 million new first patent filings in 2015, added to the 15 million that are still alive, the information is here, it is flooding in from every direction and increasing more than ever. Maps are becoming an essential way to simplify the data analysis and help the decision-making process.

Considering a topic, for which more than 10,000 patent families have been identified, means the analysis will be at a macro level. Large trends regarding the applications and use of the technologies are more likely to come up with for this size of collection. Often the topic covers multi-application concepts like a type of material (Polyurethane), a function (autonomy, renewable energy…) or a process (additive manufacturing).

Considering a topic with 1,000–10,000 patent families is more common, the meso-level. This means you have narrowed the topic to a market with a use or a new technology. It is ideal to get an overview of the competition and track their latest technology trends, pinpointing the emergence of new process or products.

When it comes to less than 1,000 patent families, the analysis can become finer, subtler and encompass more operational strategies; this is the micro level. Disruptive solutions, innovation pathways and freedom to operate can be outlined here.

Let’s use the artificial intelligence domain as an example. Without getting into a philosophical debate we can consider artificial intelligence as the capacity for a software or computer-controlled system to perform tasks generally associated with human beings. To define the scope of this domain we have used specific patent classifications (G06N1 from the International Patent Classification, Class 7062 from the US patent classification) and looked for compelling keywords (neural networks, machine learning, fuzzy logic…) (Fig. 2).

---

1 Computer systems based on specific computational models.
2 Data processing: artificial intelligence.
At a macro level, we obtain some straightforward information about the existing computing approaches to make a machine intelligent. You can see a picture of the main innovation pathways, in a past-oriented study this is very useful when you have just jumped into a new domain (i.e. you would like to be informed about what is AI). You want to “feel” the market without being buried in details (Fig. 3).

Figure 2. Macro-level global AI domain (46,300 patent families)

Figure 3. Meso-level AI applied to medical devices (2,400 patent families)
For the meso-level, we have focused on a specific market, medical devices. Immediately we notice a difference in the potential insights available in the chart. Uses, applications and technical solutions have all emerged. The main use of AI is for diagnostics using machine learning and Random Forest approaches. But electroencephalography (EEG) applications are ramping up, as is brain activity monitoring.

**Get a map from the data**

One of the most popular and instinctive moves, once we have a dataset, is to logically organize the information from the dataset. Deriving high-quality information from text is called Text Mining, often a text engine will scan through the patent collection and automatically determine the distinguishing words or “concepts” within the collection [8].

This is based on statistical measurements of word distribution, frequency, and co-occurrence with other words. Once you extract these concepts, you have the possibility to measure the distance between each document based on their shared concepts, and from there start cluster computing to group the closest documents together. Adding labels that represent each of the clusters and your map is born (Fig. 4).

The advantages are as plain as day. You get the big picture; it is visually appealing and easy to understand. You can talk about it, see relationships, and it is a great tool to share and communicate on. By using it dynamically you can literally explore the dataset, you may select the isolated documents trying to understand why they do not connect with the others, and what kind of new concepts they describe. The more crowded is an area, the more similar the topics of the documents are. If you only analyze patents, each island in the map becomes a battle of IP rights. In our use case, the neural networks are intensively protected to diagnose diseases from physiological signals. Some freedom to operate considerations could even be outlined once you set up a dataset including patents, product specifications and articles.

It is clearly seen from the above chart that for both medical imaging devices and medical diagnoses made on the basis of physiological parameters, some products have to be concerned by the important number of patents very close. This kind of map could be the starting point of uncovering potential infringers, finding monetization targets or assessing interdependence with competitors.

As ideal as these maps seem, the method can suffer from selective perception, only looking for concepts that are expressed in the text and standing out the mainstream pathways. The risk is that what you pay attention to will be determined by what you expect to see. Therefore, this angle of view could be complemented by a more rationalized approach.

**From mapping data to a taxonomy**

Patent engineers are tempted to use vague terms or uncommon words to stick to a scientific approach of an invention. But who has never wanted to organize data according to a familiar vocabulary? Does it make more sense to display separate clusters for multi-modal mobile entity and portable devices, or to create a cluster gathering all under connected objects?
Chevalier B. Leverage Patent Analytics to Achieve Business-Oriented Objectives...

127
Setting up a taxonomy, based on the knowledge of experts and previous data mining work, allows us to search for non-obvious solutions and objectively focus our attention on all the segments (see Fig. 5).

| Approach                      | Medical Specialty                        | Services                          |
|-------------------------------|------------------------------------------|-----------------------------------|
| Artificial Neural Networks    | Anesthesiology                           | Diagnosis                         |
| Support Vector machines       | Cardiovascular                           | Prediction                        |
| Bayesian Networks             | Chemistry                                 | Data mining                       |
| Expert System                 | Dental                                    | Automation/Autonomy               |
| Random Forest                 | Ear, Nose, and Throat                    | Creativity                        |
| Decision Tree                 | Gastroenterology and Urology             | Recognition (facial, visual, speech...) |
| Deep Learning                 | General Hospital                          |                                   |
| Reinforcement Learning        | Hematology                               |                                   |
| Genetic Algorithms            | Immunology                               |                                   |
|                               | Microbiology                             |                                   |
|                               | Neurology                                |                                   |
|                               | Obstetrical and Gynecological            |                                   |
|                               | Ophthalmic                               |                                   |
|                               | Orthopedic                               |                                   |
|                               | Pathology                                |                                   |
|                               | Radiology                                |                                   |
|                               | Toxicology                               |                                   |

Figure 5. Taxonomy at 3 levels for AI applied to medical devices

Supervised machine learning methods help to create multilabel classifier. More and more organizations, corporate or public entities like USPTO or EPO, are using a bot\(^3\) to classify patents according to a taxonomy. Using samples, they teach a bot to recognize specific markers that will lead to assigning labels or segments. Therefore, the task becomes very fast in regard of the manual review.

Otherwise, this analytic approach seeks to reduce a domain to its basic elements in order to study in detail and understand the types of relationships that exist between them (Fig. 6).

To be powerful, the taxonomy must embrace the different features of an invention, from its nature (component, process, material...) to its market. In our use case, it makes sense to make up a list of applications to translate the benefits and effects of artificial intelligence into some uses. It also makes sense to look for solutions for any medical specialties as well. Only then you will be able to showcase not only crowded areas but the empty zones too. In our cross-correlations matrix, there is no existing patent for dental applications. The next questions would be, is there any interest in using an AI for dental applications? Are there insurmountable technical issues?

The map will lead us to explore new innovation pathways and ask questions about the market. But above all, it will provide a solid foundation for any strategical analysis to come.

**Measuring maturity**

Because a landscape gives the opportunity to compile all data on a specific topic, it must deliver an accurate insight about the maturity. Every

\(^3\) An autonomous program on a network (especially the Internet) which can interact with systems or users.
market, product or technology has life cycle stages, correlated but with different paces, profiles and lifespans. Because patents are only one of the possible outputs of R&D efforts, they’re also strongly related to the protection of products, thus they are a formidable link between R&D and the business. From the analysis of some well identified metrics, we will determine the actions associated to the topic’s maturity.

**Growth rate**

The first metric is the most obvious and the simplest to obtain. Investigating the evolution of the patent filings over time identifies the trend and measures its strength. The activity may be growing, declining or remaining flat, each of these trends could be assigned with a grade (moderate, intense, strong).

Since 2010, the average annual growth rate of patent applications worldwide has been at 7.7 %. Any growth rate above this average could be considered as intense, but some rates can reach some impressive peaks: drones (30 %), China (45 %).

As a rule, patent filing activity reacts to the viability of a new technology and is visible before the product launch. A strong growth is symbolic of an emergent or fast-developing market where a new field of innovation is beginning.

Going back to our example, the use of artificial intelligence methods or approach with medical devices has ramped up strongly since 2012. The activity is going through a strong growth of 33 % annually and is easily identified as an emergent domain because of its really low number of patent applications before 2010. More patents have been published for the last 5 years than during the 15 prior.

![Figure 6. Cross-correlations between medical specialties and services](image-url)
IP events lifecycle

Every market, product or technology has life cycle stages and Gartner, an information technology (IT) research and consultancy company, has even developed a graphical representation, the hype cycle, to illustrate how a technology goes through from conception to maturity and widespread adoption. For each stage, a specific behavior related to patent activities can be predicted (see Fig. 7).

![Figure 7. Patent life cycle according to the Hype Cycle](image)

Patent filings follow a technology outbreak at the very beginning, scattered at first and then with sustainable growth when the products are launched and commercially viable. When the peak of expectations is at its highest, the pioneering patent rights are starting to whet the appetite of the first investors, first transactions are made and will intensify continually. The slope of enlightenment is described as the moment where the market knows how the technology can benefit the enterprise and becomes more widely understood, and, while new patents are filed, this period represents companies’ awareness that the technology they lack is not in their laboratories. The first blocking patent portfolios emerge, and licenses become common. At the end, the technology is mainstream, to sell more, you need the competitors to sell less. You start to watch for counterfeiters, attempt to slow down any challengers and attack all new entrants.

With the most recent patent database providers, access to litigation, opposition or technology transfer data has become very easy. From your collection of documents, you have the possibility to benchmark these data points with the technical area references (see Fig. 8).

Having a look at our use case, figures show that this is not an aggressive domain, far from it! But the first licenses and litigations have appeared, highlighting a better understanding of the benefits of the technology and the beginning of commercial viability [9].
Figure 8. IP events (litigation, licensing, opposition) for the area of AI applied to medical devices

Number applicants by number of patents through years

The last metric has been practiced by the academics for years now. Sometimes called the innovation cycle of a technology, it is based on the concept that a vast amount of information can be gathered about the cycle and maturity of a given technical field through the analysis of how diverse and active the applicants are: a succession of development, maturation and recovery periods provide rhythm to the life cycle of a technology. In the developing period of a technology, the number of applicants and applications rapidly increase, and research interest and development actively take place. Inversely, when the technology market is shrinking, number of applicants and applications both decrease (see Fig. 9).

The use of artificial intelligence in medical devices is maturing right now. After a long period (2011–2015) of development to understand how useful this technology could be, the activity is entering into a phase where just the specialists and the survivors (those who found the correct solutions) continue applying for patents.

Merging the different inferences from the analysis of the three metrics provides us with a clear and objective opinion about the technological maturity of our topic or innovation pathway. Incorporating market information, like mergers and acquisitions, product launches, R&D spending, will help to illustrate the final conclusion and make sure we do not miss any important recent events. We have shown you how to find the main technological pathways, from a map, and determine if the market is experiencing a phase of technological consolidation or, quite the opposite, a turmoil of new disruptive solutions. Our last
step will touch the essence of a landscape, interpreting any weak signals
to anticipate the future success.

**Weak signals & trends**

At its best, data is what is known in economics as a “lagging indica-
tor,” a rearview-mirror glimpse of a previous reality, which does not make
it the best prediction tool at first sight. But if you know how to amplify
interesting information and keep seeking new data to confront the actual
reality, you will encourage a salutary constructive conflict with your way
of thinking, leading down the road to success for your business.

**What is a weak signal?**

You can define a weak signal as a disconnected, insulated piece of
information that at first appears to be just background noise. But it can
become a compelling pattern by analyzing it with other pieces of informa-
tion and the starting point of a new mainstream innovation [10].

**How do you find them?**

You need different approaches, a little sweat and skills for synthesis
and identifying relationships. Here are our methods:

– focusing on recent patents (age < 2 years); You will find the new
technical issues and pinpoint the next market shifts.

– focusing on deserted areas on the map; Do these patents and articles
describe avant-gardist solutions? Why hasn’t anyone followed those solu-
tions? Cost, flaws, complexity? What is not displayed on the map?

– analyzing new patent classes/new concepts appearances these last
years; What are the inventions linked to these apparitions? How have
these patent classes been used before for this topic? What has changed?
– analyzing “unusual” US proceedings (accelerated examinations, requests for reexamination…); These practices are pushed by business incentives.
– analyzing publications recently cited by patents; Popular scientific publications cited by patents could be the starting point of disruptive innovations.

**What do you do with that?**

Based on our use case, we have looked for patents in deserted areas and including new concept appearances. A dozen very recent patents, all filed after 2014, have been pinpointed, describing the use of an avatar to assess the emotional state of a patient or monitor their biometric feedback. The principle of developing patient avatars could be associated with improvement of health outcomes, simulating clinical trials and for patient-to-patient peer interactions to improve personalized self-care (see Fig. 10).

For each weak signal you need to build a scenario and compare it to your business roadmap. For example, do you cover this R&D area? Will your existing products be impacted by such computational avatars? Could one of your services be cannibalized by this technology?

You will need to talk to your customers and suppliers about the scenario, comparing your diverse hypotheses to the reality. It is at this point that your perspective will become broad enough to avoid the biases, to see what is really going on and provide a foundation for choosing a path. You will know the arguments to choose between an internal or external innovation pathways.

---

**Figure 10. Isolated patents & Potential white spaces**
Regarding our use case, we have discovered that some solutions have found success, we are entering a maturation period where the smartest will get benefits from the growth and the viability of the products. There is no time to start an innovation project, you need to keep up and monitor your freedom to operate status. The next moves may be to align your existing IP portfolio with your products and order freedom to operate studies. Your innovation must be incremental and focus on the viable technologies and start looking for external technologies.

Conclusion

What worked in one context, at a given time, will not always guarantee a bright and sustainable future. That is why, since 2000, half of the companies in the Fortune 500 have disappeared. You need to anticipate, multiply the hypotheses, seek the conflict and pay attention to the weak signals.

A landscape of patents and technological documents is the perfect tool for that. The mapping process will help you to see your world in a different way than what you have in your mental map. Your cognitive and emotional biases will be challenged and reduced to make way for a more objective and relevant interpretation of the weak signals. What we have found for the topic of AI-based system in medical devices is very promising with real innovation pathways outlined. But we are certain that all these results will be without comparison with what will be produced with the new deep learning methods used to classify and map patents.

In this fast-moving world, you cannot afford to miss the next technological revolution or the next incremental innovation of your competitor. The difference will be determined by gathering the data and interpreting its meaning.

References

1. Bowen Y, Jianxi L. Measuring Technological Distance For Patent Mapping. Journal of the Association for Information Science and Technology. 2016; 68 (2):423-437. DOI: https://doi.org/10.1002/asi.23664
2. Aristodemou L, Tietze F, Athanassopoulou N, Minshall T. Exploring the Future of Patent Analytics: A Technology Roadmapping approach. Centre for Technology Management Working Paper Series. 2017. 5:1-10. DOI: https://doi.org/10.17863/CAM.13967
3. Kukolj D, Ostojic G, Stankovski S, Nemet S. Technology Status Visualisation Using Patent Analytics: Multi-Compartment Refrigerators Case. Journal of Mechatronics, Automation and Identification Technology. 2019; 4 (1):1-8. Available at: http://jmait.org/httpjmait-orgvolume-4-issue-1-2019-1-8/ (accessed: 25.02.2020)
4. Suominen A, Toivanen H. Unsupervised Learning-Based Patent Landscapes Using Full-Text Patent Data. Portland International Confer-
Chevalier B. Leverage Patent Analytics to Achieve Business-Oriented Objectives...

5. Rattinger A, Le Goff J-M, Meersman R, Guetl C. Semantic and Topological Patent Graphs: Analysis of Retrieval and Community Structure. 5th International Conference on Social Networks Analysis, Management and Security, SNAMS. 2018; 51-58. DOI: https://doi.org/10.1109/SNAMS.2018.8554761

6. Yeap T, Loo GH, Pang S. Computational Patent Mapping: Intelligent Agents for Nanotechnology. Proceedings International Conference on MEMS, NANO and Smart Systems. 2003; 274-278. DOI: https://doi.org/10.1109/ICMENS.2003.1222008

7. Lee CH, Yang HC, Li YJ. Development of A Multilingual Text Mining Approach for Knowledge Discovery in Patents. IEEE International Conference on Systems, Man and Cybernetics. 2009; 2334-2338. DOI: https://doi.org/10.1109/ICSMC.2009.5345953

8. Fattori M, Pedrazzi G, Turra R. Text mining applied to patent mapping: a practical business case. World Patent Information. 2003; 25(4):335-342. DOI: https://doi.org/10.1016/S0172-2190(03)00113-3

9. Li Z, Atherton M, Harrison D. Identifying Patent Conflicts: Triz-Led Patent Mapping. World Patent Information Volume. 2014; 39:11-23. DOI: https://doi.org/10.1016/j.wpi.2014.07.002

10. Guo M, Yuan H, Qian Y. A New Method for Rare Feature Extraction in Patent Documents. 13th International Conference on Service Systems and Service Management (ICSSSM). 2016; 1-6. DOI: https://doi.org/10.1109/ICSSSM.2016.7538569

Submitted: 27.02.2020
Дата поступления: 27.07.2020

Information about the author
Benoit Chevalier, Engineer Electronics and Telecoms, Senior Consultant at Questel (1 boulevard de la Madeleine, Paris 75001, France), co-founder of Avenium Consulting (now Questel Consulting), Member of LES France (Licensing Executive Society), ORCID: https://orcid.org/0000-0001-8357-1631.

Information об авторе
Шевалье Бёнуа, инженер телекоммуникаций и связей, старший консультант, Questel (75001, Франция, г. Париж, Бульвар Мадлен, д. 1), сооснователь Avenium Consulting (сейчас – Questel Consulting), член французского регионального отделения Международного лицензионного общества (Licensing Executives Society, LES France), ORCID: https://orcid.org/0000-0001-8357-1631.