Machine Translation and Disclosure of Patent Information

Aline A. Larroyed

Published online: 24 July 2018
© The Author(s) 2018

Abstract Machine translation represents a revolution in the patent system. It facilitates access to a vast amount of patent information on a global level during the different stages of patent prosecution, especially in searches for prior art documents. Moreover, disclosure represents one of the main arguments to justify the existence of patent rights, but how much of this information is actually disclosed by machine translation? The present research aims to investigate the level of disclosure of patent information made possible by machine translation and its implications for patent law. The guiding research question is: How and to what extent is patent content actually disclosed by machine translation? In order to reach the results, a manual method of translation quality assessment was used to analyse a sample of 100 patents, combined in different pairs of languages. Machine translated patent texts were also submitted for evaluation by researchers skilled in the art. The research is also based on a review of legal and theoretical sources. The results proved that, although improvements in patent machine translation are still necessary, most of the content of the patent is disclosed through it, with the average disclosure being almost 80%. As the concept of disclosure is a central one to patent law, it is hoped that this research will contribute to a better understanding of the current role of machine translation in patent law.

Keywords Machine translation · Disclosure · Patent content · Person skilled in the art · Patent law

A. A. Larroyed (✉)
Visiting Researcher Max Planck Institute for Innovation and Competition (Munich, Germany); PhD
Candidate Maastricht University, Maastricht, The Netherlands
e-mail: Aline.Azevedo-Larroyed@ip.mpg.de

1 Introduction

The configuration of contemporary techno-scientific development has moved patent law to a global level and to the centre of the discussion regarding scientific knowledge and its respective technological applications. These developments have led to new levels of communication and information exchange, which, due to the size and features of the patent system, are currently made possible by the use of translations made by computer software, so-called machine translation.¹

Patent machine translation has become an increasingly sophisticated communication tool as a result of continuous efforts in reaching accuracy in patent expert systems. These developments have lowered language-related barriers and permitted the disclosure of patent information to be more effective. Therefore, machine translation represents an important improvement related to the main role of patent rights: the disclosure of the content of inventions in exchange to the grant of exclusive rights to the patentee.

Moreover, patents are seen as a prerequisite for competitive markets, as patent rights allow for the impartation of knowledge, the prevention of imitation of the imparted content and contribute to fair market behaviour.² In order to fulfil this function, patents should be granted only if they disclose the invention “clearly and completely to the person skilled in the art”.³ Machine translation has also changed the reality of patent examination and patent prosecution, giving examiners, attorneys and interested parties access to a vast amount of information.

Consequently, the research question guiding the present investigation can be considered as a central one to patent law in its current global configuration: How and to what extent is patent content actually disclosed by machine translation? The evolution of the patent system has led to a tendency toward integration and harmonisation, which demands fast and efficient communication. In terms of international patent law, this new configuration has been enabled by different instruments, such as international agreements, cross-border institutions and cooperation between patent offices.⁴

Some of these initiatives are particularly noteworthy. The European Patent Convention (EPC), established in 1973, is an example of a large step towards integration and harmonisation. The EPC made possible the creation of the European Patent Office and the European Patent, thereby simplifying the procedures (including language and translation) for patent applications in Europe. The regulations to implement enhanced cooperation in the creation of a Unitary Patent and a Unitary Court represent an even more significant step in this

¹ Cyril et. al. (2010).

² WIPO, “Competition and Patents”, available at: <http://www.wipo.int/patent-law/en/developments/competition.html> (accessed 4 July 2018).

³ Article 83 of the European Patent Convention (EPC), 16th edn. June 2016, available at: [http://documents.epo.org/projects/babylon/eponet.nsf/0/1B8F3A8943AC2EF3C12581E800591D5A/\\$File/EPC_16th_edition_2016_en.pdf](http://documents.epo.org/projects/babylon/eponet.nsf/0/1B8F3A8943AC2EF3C12581E800591D5A/$File/EPC_16th_edition_2016_en.pdf) (accessed 4 July 2018).

⁴ Nam and Barnett (2011), pp. 1471–1485.

direction, which is combined with the advent of more efficient tools for automatic translation.⁵

The EPO's "Guidelines for Examination" include numerous references related to translation as well as a section dedicated to machine translation. The document clearly outlines the relevance of machine translation to patent examination by stating that, "in order to overcome the language barrier constituted by a document in an unfamiliar non-official language, it might be appropriate for the examiner to rely on a machine translation of said document".⁶

Furthermore, the Guidelines are assertive in recognising the use of machine translation as a legitimate means to access and understand prior art: "A translation has to serve the purpose of rendering the meaning of the text in a familiar language [...] Therefore, mere grammatical or syntactical errors which have no impact on the possibility of understanding the content do not hinder its qualification as a translation".⁷

However, language differences and translation still represent an obstacle for establishing agreements that cover multilingual communities. Although machine translation can be observed as the most promising tool to diminish these burdens by fulfilling the demands of a growing volume of multilingual communication, it continues to provoke distrust and controversies. The research presented in this article aims to provide some answers to these controversies by clarifying how much information machine translation actually discloses and how far it still has to go in order to achieve a level of excellence in carrying out this function.

The methods used in reaching the conclusions involve the empirical analysis of primary sources, due to the lack of previous studies on the topic. These methods are thoroughly explained in Sect. 2. They include: working with different pairs of languages, comparing patent content from different search platforms, using a manual technique to measure the quality of the translation, and comparing the initial results with the evaluation of persons skilled in the art represented by scientific experts in the technological field of the chosen sample. The subsequent sections discuss the results of the investigation and present recommendations for future studies that can contribute to a better understanding of the disclosure of patent information through machine translation. Section 2 also presents the methods used to reach the results. This section aims to introduce the details of the methodology. Due to the level of novelty of this study, it was necessary to combine methods and techniques in order to reach the aimed results. Section 3 examines the concept of disclosure within the scope of the European patent. Section 4 discusses patent writing associated with patent translation and disclosure. Section 5 presents the results of the analysis, concluding that machine translation not only discloses patent

⁵ The whole set of regulations on the Unitary System is available at Bristows UPC [website]: <https://www.bristowsupc.com/legislation-and-documents/> (accessed 4 July 2018). See Council Regulation (EU) No. 1260/2012 of 17, December 2012, implementing enhanced cooperation in the area of the creation of unitary patent protection with regard to the applicable translation arrangements, available at: http://www.wipo.int/wipolex/en/text.jsp?file_id=308371 (accessed 4 July 2018).

⁶ EPO, "Guidelines for Examination", Sec. 4.1, available at: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_iv_4_1.htm (accessed 4 July 2018).

⁷ *Ibid.*

information, but may be offered as a solution to the controversies of patent content disclosure and be seen as an important tool in agreements and initiatives that will bring about more integration and harmonisation in the patent system. Section 6 provides a summary of the conclusions and recommendations for further studies.

2 Methods and Techniques Used to Reach the Results

The research is primarily based on empirical analysis and on the examination and combination of data collected from primary sources due to the novelty of the research question. Defining and narrowing a sample of patents represented an arduous task, considering the size and variety of the international patent system. The first step in narrowing the sample was to focus on green technologies, as prioritised by the Institute where the study was developed – the Institute for Globalization and International Regulation (IGIR) at the Faculty of Law of the University of Maastricht, the Netherlands. The second stage was to combine the information disclosed by the patents with their (machine) translated texts and with the analysis of collaborators representing the ideas of persons skilled in the art, e.g. scientists belonging to the same technology field.⁸ The analysis was also based on a review of literature related to the concept of disclosure of patent information.

The main sources consulted for the literature review encompass legal and theoretical sources. The main legal references consulted were the agreements and regulations associated with the European patent, especially the EPC, the “Guidelines for Examination” of the EPO and the set of Regulations of the European Unitary System.⁹ The choice was made due to the history of the European patent system with patent translation and the fact that the European system is considered as a benchmark in terms of integrating and harmonising national patent laws. The European regulations also provide clear legal paradigms to the concepts on which this research is based, especially the concept of disclosure.¹⁰ The theoretical texts were chosen among the main references relating to the idea of disclosure in patent law, to patent writing and patent translation.

In order to acquire complete information regarding the sample, it was decided to work with three databases in a complementary manner, two of them being paid databases, Patsnap and DartsIP, and one of them, Espacenet, being the free database offered by the EPO. The three chosen databases offer different types of filters, focusing on different sorts of information about the patents – all available through machine translation.¹¹ The analysis of the sample of 100 patents, complemented

⁸ According to the “Guidelines for Examination” of the EPO, a person skilled in the art “should be presumed to be a skilled practitioner in the relevant field of technology, who is possessed of average knowledge and ability and is aware of what was common general knowledge in the art at the relevant date”. See the complete definition at: http://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_vii_3.htm (accessed 4 July 2018).

⁹ See *supra* notes 3, 5 and 6.

¹⁰ See *supra* notes 3, 5 and 6.

¹¹ The homepages of the platforms are available at: <http://www.patsnap.com/>; <http://www.darts-ip.com/>; <http://www.epo.org/searching-for-patents/technical/espacenet.html#tab1> (accessed 4 July 2018).

with 10 patents and supported by native speakers of the chosen languages, was conducted manually and lasted one year, as it required a word-by-word, sentence-by-sentence, comparison. It was important to combine the information from the three platforms in order to obtain a reliable sample.

Patsnap is an analytic platform that focuses on the competitive landscape of patents. It provides tools which allow companies to plan their IP strategy according to the contexts that surround a technology. For the present research, it was the most suitable platform for understanding the scenario of green technologies and for selecting a valid sample due to its economic relevance and geographical/linguistic reach.¹²

Darts-IP is an expert system that focuses on industrial-property-related case law. It makes possible searches by points of law, parties, patents or courts. Its free text search tool also supports different searches in different languages (separately or simultaneously). In the case of this research, it was useful for accessing the most important legal issues involving the sample, providing an idea of the global legal scenario that surrounds the selected patents. It also helped to sort out only those patents which were already granted.¹³

Espacenet, the free platform developed by the EPO in 1998, is a very complete platform in terms of quantity of data and variety of languages that can be machine translated. It provides data on more than 90 million patents from around the world. The search engine offers a tool called SmartSearch, which allows a composed search using a subset of Contextual Query Language.¹⁴ For this research, one service considered especially useful was “Patent Translate”, which was used as the main reference for the assessment of the quality of patent machine translation.¹⁵

The sample of green patents was narrowed while it became possible to get more familiar with the databases and the potential types of technologies that would compose the sample. The criteria used to choose the type of technology were its economic importance, the time period of publication (limited to patents published between 2012 to 2015), and the possibility of finding a reasonable amount of patents with international relevance. After testing some possible combinations of patents, the sample was restricted to wind power and solar energy patents, with the understanding that these would provide enough information to the aspects under analysis and would also be representative in terms of economic, social and geographic relevance (Table 1).

By combining the information contained in the three platforms mentioned above, it was decided to carry out the analysis in two steps. The first step aimed to assess how much of the patent’s information is disclosed by a machine translation through a manual method, with the support of native speakers of some of the chosen languages.¹⁶ The second step aimed to confirm the results, by submitting machine

¹² *Supra* note 11.

¹³ *Supra* note 11.

¹⁴ EPO Espacenet, available at: <https://www.epo.org/searching-for-patents/technical/espacenet.html#tab-1> (accessed 4 July 2018).

¹⁵ *Supra* note 11.

¹⁶ The choice for a manual method was based on research related to reliability of machine translation assessment. *See* White (2003), pp. 211–244.

Table 1 Criteria for selecting the sample. Source: based on the information collected through the patent databases. (See *supra* note 11.)

Economic importance
Proximity in time – from 2013 to 2015
Legal situation – only granted patents were selected (with this criterion, it was intended to assure that the patent text had a minimum quality to fulfil the requirements of the analysis, as granted patents are the ones which actually correspond to new inventions)

translated patent documents to the analysis of persons skilled in the art, working in that field of technology.

To perform the first step of the analysis, it was important to initially evaluate the possibilities of working with different sizes of the sample related to clean energies, consisting of wind power and solar energy patents. For this purpose, a sample size calculator was used on the average number of these patent applications per year given by Patsnap for the period 2013, 2014 and 2015 (considering only granted patents).¹⁷ As the average number of patent results was 6,523 patents, and considering a margin of error of 10%, a confidence level of 90% and a response distribution of 50%, the minimum recommended sample size was 67 patents. In order to work with a round number, 100 patents were selected among the applications made between 2013 and 2015.

The set of languages chosen for the analysis included English, French, German, Italian, Spanish, Chinese and Portuguese. These languages were chosen according to the availability of native speakers to provide support to the analysis. Portuguese was the main reference as a target language in the process of analysis. The inclusion of Chinese is justified by the noticeable percentage of patents in Chinese among the applications in the selected sample (Table 2).

After narrowing the sample and defining its characteristics, it was necessary to select the method to assess the quality of the machine translations provided by the EPO's Patent Translate. It was decided to base this assessment on the LISA (Localization Industry Standards Association) Quality Assessment (LISA-QA) default model – a system created for managing the quality of translations certificated by ISO 9001. According to this metric, errors are regarded as “Minor”, “Major” or “Critical” and separated into six categories.¹⁸

For each category, the number referring to the error is multiplied by a weighting figure, as depicted by the table below. For example, if four minor mistranslation errors are identified, they generate a score of 4; but if a segment of a text contains 2 critical translation problems, it should be written twice 10 on the column

¹⁷ The Raosoft sample size calculator is available at: <http://www.raosoft.com/samplesize.html> (accessed 4 July 2018).

¹⁸ See details about the metric in the Localization Industry Standards Association's official website: <http://web.archive.org/web/20110101195746/http://www.lisa.org/LISA-QA-Model-3-1.124.0.html> (accessed 4 July 2018). See also “LISA Best Practice Guide – Implementing Machine Translation”, available at: http://www.mikedillinger.com/papers/MT_BPG2004.pdf (accessed 4 July 2018) and http://web.archive.org/web/20110102030235/http://www.lisa.org/globalizationinsider/2010/08/the_lisa_qa_mod.html (accessed 4 July 2018).

Table 2 Description of the sample used for the first step of the analysis. Source: the table was elaborated by this author in order to categorise the information on the sample and its characteristics.

Types of patents	Green patents – clean energy – solar energy and wind power patents
<i>First Step of the Analysis – Level of Disclosure: Description of the Sample</i>	
Search by geographical location	<i>First set:</i> patents from different parts of the world except from China and Japan aiming to find the selected languages: English, French, German, Italian and Spanish <i>Second set:</i> patents from China
Languages	<i>First set:</i> English, French, German, Italian, Spanish and Portuguese <i>Second set:</i> English (target language) and Chinese (source language)
Sample size	100 patents, chosen from an average number of 6,523 patents per year (considering a margin of error of 10%, a confidence level of 90% and a response distribution of 50%)
Platforms used	Patsnap, Espacenet and Darts-IP
Time period of the selected sample	Patent applications from 2013 to 2015

for “Critical Mistranslation”, which results in a score of 20.¹⁹ The table below describes the method used for characterising each translated document (Table 3).

Although the LISA-QA model does not establish a level for assuring excellence in translation (some companies work with 99%),²⁰ considering the type of translation object of this analysis and the possibility of counting on the support of source texts and drawings, a lower level of 70% correct (30% errors) was established as a parameter. The texts were analysed one by one and the results were registered manually in order to assure their reliability.

Other important information regarding the sample analysis is the part and size of the text of the patent selected, in order to obtain more precise results.²¹ The two parts of the patents which were analysed – the abstract and the claims – are considered crucial to the disclosure of technological information, also being representative of the content of the invention in terms of textual structure.²² It was decided not to include the detailed description of the invention or the drawing set, in order to focus on text information only, as this type of information would make the disclosure of the invention much easier.²³ Working with the selected parts, in this

¹⁹ *Ibid.*

²⁰ *Ibid.*

²¹ Section 4 of this article will provide a thorough description of the structure of a patent text and its relationship with translation and disclosure.

²² For the structure of a patent text, and the form and content of the abstract and claims of a patent, see EPC, Rules 43 and 47. See also the EPO “Guidelines for Examination” Chapter II – Content of a European patent application (other than claims). Available at: <https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/r43.html>; <https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/r47.html>; and https://www.epo.org/law-practice/legal-texts/html/guidelines/e/f_ii.htm (accessed 4 July 2018).

²³ For the sections of a patent document, see: “WIPO Patent Drafting Manual”, pp. 33–42. Available at: http://www.wipo.int/edocs/pubdocs/en/patents/867/wipo_pub_867.pdf (accessed 4 July 2018).

Table 3 LISA-QA – metric of translation errors applied to the sample. Source: based on LISA's Best Practice Guide. (See *supra* note 18.)

	Minor	Major	Critical
Doc Language	1	5	10
Mistranslation	1	5	10
Accuracy	1	5	10
Terminology	1	5	10
Language (syntax)	1	5	10
Style	1	5	10

sense, meant to choose the more stringent way, in order to achieve a higher level of confidence in the analysis.

In terms of text extension, the abstract contains a reasonable amount of textual information, which is manageable in the size defined for the sample. However, in order to achieve a homogeneous sample, it was decided to work with the first 100 words of this part of the patent text (having the flexibility to add or exclude a small number of words in order to have complete sentences).

The patent claim can be represented by a much longer text than the abstract. It is a crucial text for the disclosure of the technical information and a more technical text than the abstract in terms of its language structure, since the abstract functions as an introduction or a summary of the patent. According to the EPC, “the claims shall define the matter for which protection is sought in terms of the technical features of the invention, and the abstract shall contain a concise summary of the disclosure as contained in the description, the claims and any drawings”.²⁴

Hence, in order to achieve more homogenous and manageable text material, the size of the patent claims was chosen to be approx. 200 words, a length which was sufficient for having enough textual content for a reliable analysis. This last figure was defined after some practical work with the sample and was shown by the characteristics of the texts of the selected patents (Table 4).

For each patent, there was a form containing patent information considered as crucial (title, economic value, publication number and date, application number and date, assignee, original language), the source and target texts to be analysed, and the score table based on the LISA-QA model. An example of the score table is presented below (Table 5):

In addition to using the LISA-QA model-based analysis, it was understood that it would be relevant to assess how a person skilled in the art belonging to the research field of the sample would react to the machine translated texts, considering the question to be answered in this step of analysis concerned to what extent a machine translated text of a patent could disclose the content of the invention claimed in it.

²⁴ *Supra* note 18.

Table 4 Details of the sample size. Source: developed on the basis of the information contained in Patsnap and Raosoft. (For Patsnap, see *supra* note 11; and for Raosoft, *supra* note 17.)

	Total number	100
Number of patents	“English patents” – patents having English as their source language and Portuguese as target language	50 (30 on solar energy and 20 on wind power)
	“French patents”	10 (+ 10 for confirming the high score/few errors results). The last 10 patents were added to the sample
	Target language: Portuguese	
	“German patents”	10
	Target language: Portuguese	
	“Spanish patents”	10
	Target language: Portuguese	
	“Italian patents”	10
	Target Language: Portuguese	
	“Chinese patents”	10
	Target language: English	
Level of correction	At least 70% correct (30% of errors) to be considered as a reasonable translation.	
Size of the texts	Abstract: $\pm$ 100 words	Claims: $\pm$ 200 words (results were divided by 2)

For this purpose, it was necessary to develop another type of method and form to be shared with the collaborators. The analysis was slow and laborious. However, the researchers who cooperated did so voluntarily and with the agreement of their supervisors, allowing part of their work time to be dedicated to it. The form which was supplied had ten pages and, according to them, took from two to four hours to be analysed and answered. Below is an example of part of the form used with the researchers to assess the level of disclosure of patent information through a machine translation (Fig. 1).

The three texts shared with the researchers were chosen at random among the most relevant results displayed on Patsnap regarding the sample. Text number 1 was a (machine) translation into Portuguese of a patent written originally in German; text number 2, a translation into Portuguese of a patent written originally in English; and text number 3, a translation into Portuguese of a text written originally in Chinese.²⁵ The researchers were not informed of the original language of the texts. They only knew that they were machine translated into Portuguese.

The diagram below summarises the steps followed to conclude the analysis (Fig. 2).

²⁵ It was not possible to obtain a direct machine translation from Chinese into Portuguese using the databases. Espacenet provides a machine translation of the text from Chinese into English, and Patsnap provides human translation of all the Chinese patents into English. The English translation of Espacenet was used. It was retranslated using Google Translate into Portuguese (a way a researcher would most probably naturally follow to understand the Chinese text).

Table 5 Example of score table used to evaluate the translation of the patent texts. Source: developed on the basis of LISA's Working with the selected parts, in thBest Practice Guide. (For Putsnap, see *supra* note 11; and for Raosoft, *supra* note 17.)

TITLE: Alternating Bias Hot Carrier Solar Cells									
Value: : \$ 1,500,000			Assignee: OSTENDO TECHNOLOGIES, INC.						
Pub. N.: US20140083492A1			Pub. Date: 27 Mar 2014						
App. N.: US14/028344			App. Date: 16 Sep 2013						
Doc. Language(s):	Minor = 1		Major = 5		Critical = 10				
	Abstract	Claims	Abstract	Claims	Abstract	Claims			
Mistranslation				1		2			4
Accuracy	1		1						
Terminology									
Language	3	7		3		1			
Style									
Total score	Abstract	39 (100 words)		Claims	67 (200 words)				

Answer to both statements and to the questions below for each of the following patent texts. Please do not refer to the original texts, by consulting solely the contents of this form. Each text contains abstract and claims of the patent.

Statement 1: Is the translated text clear in describing the content of the invention?	Totally disagree	Disagree	Neutral	Agree	Totally agree
Text 1					
Text 2					
Text 3					

Statement 2: does the translated text allow the disclosure of the content of the invention?	Totally disagree	Disagree	Neutral	Agree	Totally agree
Text 1					
Text 2					
Text 3					

Which one of the texts was clearer in describing the invention?

If you judge necessary, justify your answer.

Figure 1 Part of the form shared with collaborators for assessing the level of disclosure of patent information through machine translation. Source: the form was elaborated by this author based on qualitative analysis methods and techniques. (See Krippendorff (2004). See also Berg (2001), p. 255.)

3 The Concept of Disclosure Under the Perspective of the European Patent

Imparting knowledge is one of the main *raison d'être* of the patent system, the disclosure of technological content being one of its fundamental bases. Nevertheless, the concept of disclosure still inspires discussion concerning its role and legal implications for the process of patent bargaining, and with respect to fostering innovation is subject to questions and doubts.²⁶ Disclosure can basically be defined as the public dissemination of the content of an invention or, in more technical terms, the clear revelation of the inventive matter of a new technology.²⁷ In the current scenario of the patent system, disclosure is aligned with the possibility of accessing patent documents through machine translation, as acknowledged by the EPO's "Guidelines for Examination":

A general statement that machine translations as such cannot be trusted is not sufficient to invalidate the probatory value of the translation. If a party objects to the use of a specific machine translation, that party bears the burden of adducing evidence (in the form of, for instance, an improved translation of the whole or salient parts of the document) showing the extent to which the

²⁶ For the relationship between patent rights and the role of the patent system for fostering innovation, see Landes and Posner (2003).

²⁷ Gibbs and DeMatteis (2003).

Figure 2 Diagram representing the steps followed in the analysis. Source: the form was elaborated by this author based on qualitative analysis methods and techniques. It was not possible to obtain a direct machine translation from Chinese into Portuguese using the databases. Espacenet provides a machine translation of the text from Chinese into English, and Patsnap provides human translation of all the Chinese patents into English. The English translation of Espacenet was used. It was retranslated using Google Translate into Portuguese (a way a researcher would most probably naturally follow to understand the Chinese text).

quality of the machine translation is defective and should therefore not be relied upon.²⁸

Conceptualising disclosure is not simple, due to the fact that its definition depends on national law and a legal understanding of related concepts. Article 83 of the EPC defines disclosure as a pre-requisite of the European patent application, which “shall disclose the invention in a manner sufficiently clear and complete for it to be carried out by a person skilled in the art”.²⁹

The definition of a “person skilled in the art” finds many references in the case law of the Boards of Appeal of the EPO. The definition of a person skilled in the art in the field of biotechnology, for example, determines that “his attitude is considered to be conservative”. Further, case law provides that, in genetic engineering, a person skilled in the art should not be a “Nobel prize winner”. In T 500/91, the average “skilled person” can also be considered as a team of specialists, and in T 500/91, the person skilled in the art is considered as not being someone “engaged in creative thinking”.³⁰

²⁸ EPO’s “Guidelines for Examination”, Part G, Chapter IV, 4, 4.1, available at: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_iv_4_1.htm (accessed 4 July 2018).

²⁹ Article 83 of the EPC. See *supra* note 3.

³⁰ See T 455/91, T 500/91, EPO’s “Case Law of the Boards of Appeal”, Sec. I.D., 8, 8.1, available at: http://www.epo.org/law-practice/legal-texts/html/caselaw/2016/e/cldr_i_d_8_1_3.htm (accessed 4 July 2018).

The concept of disclosure appears many times in the EPC and is also associated with other legal concepts, such as “sufficiency of disclosure”, “inventive step” and “novelty”. Among these concepts, a central one for the present research refers to “sufficiency of disclosure”, which can be defined as a legal requirement, according to which a patent application must disclose its invention “in a detailed description of at least one way of carrying out the invention”, which should be understandable to a person skilled in the art.³¹

On the other hand, the concepts of inventive step and novelty are also related to the notion of disclosure. The inventive step refers to the requirement that the invention should be “non-obvious to the person skilled in the art”.³² Therefore, the way in which the content of the invention is disclosed may affect the way in which the inventive step is examined, thus producing different results. Further, language and translation may be determinant on the decision of patent examiners in considering the inventive step when comparing the patent document with prior art. In order to avoid a subjective analysis of the inventive step, a problem and solution approach

[...] requires that the invention be disclosed in such terms that the technical problem (even if not expressly stated as such) and its solution can be understood. Problem and solution are thus component parts of any technical invention. The problem and solution approach was primarily developed to ensure objective assessment of inventive step and avoid ex post facto analysis of the prior art.³³

The Guidelines for Examination dedicate a chapter to detailing the analysis of the inventive step, showing its importance to patent examination. In examining the inventive step, some aspects and conditions should be considered by the examiner. The most significant of them are associated with the state of the art, the date of filing, the person skilled in the art and the obviousness of the invention.³⁴

The examiner should consider whether the invention would have been obvious to the person skilled in the art before the priority “to arrive at something falling within the terms of the claim”.³⁵ In other words, the inventive step is about the non-obviousness of what is disclosed in the patent claims compared to what was disclosed before its date of filing. In this regard, analysing the inventive step requires comparing texts belonging to global, multilingual platforms, a task that the examiners fulfil through prior art searches. Therefore, improving the accuracy of the information obtained through machine translation is essential to improving the quality of the examination related to the scope of inventions available through disclosure.

³¹ EPO, official website, available at: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/f_iii_1.htm (accessed 4 July 2018).

³² Article 56 of the EPC.

³³ EPO’s “Case Law of the Boards of Appeal”, Sec. I.D.2, available at: http://www.epo.org/law-practice/legal-texts/html/caselaw/2016/e/clr_i_d_2.htm (accessed 4 July 2018).

³⁴ Chapter VII (1, 2, 3, 4, 5) of the EPO’s “Guidelines for Examination”, available at: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_vii.htm (accessed 4 July 2018).

³⁵ EPO’s “Guidelines for Examination”, Chapter VII, 4.

The same is valid for the criterion of novelty, the examination of which requires prior art search and, very frequently, the analysis of foreign texts through machine translation. According to the EPC, “an invention shall be considered to be new if it does not form part of the state of the art”. Furthermore, the state of the art “shall be held to comprise everything made available to the public by means of a written or oral description, by use, or in any other way, before the date of filing of the European patent application”.³⁶ As it is possible to infer from Art. 54, the EPC has a broad interpretation of the concept of disclosure related to prior art, acknowledging even oral descriptions that could have “made the invention available, in other words, disclosed, to the public”.³⁷

In addition to being connected to the criteria of patentability, disclosure is related to language and translation, being associated with the textual framework of the patent and with the way it is interpreted during the process of examination and in further legal stages. Fromer describes the patent text as being divided into two highly relevant layers: the legal layer and the technical layer. The author observes that the attention given to the legal layer should not jeopardise the technical layer. In this respect, patent writing should comply with disclosure requirements in a broad sense. Whilst the scope of the legal layer is related to the extension and limits of the patent rights, the technical layer is directly associated with the central premise of patents in fostering innovation.³⁸

In this regard, the EPC is rigorous in determining rules over language and translation during the entire process of patent application. According to the EPC, the EPO has three official languages, English, French and German, in which all of its proceedings are managed. A European patent application shall be filed in one of the official languages or, if filed in any other language, translated into one of the official languages in accordance with the Implementing Regulations, within two months of the filing.³⁹

The Office does not conduct court cases in a strict sense, unless the court-like proceedings before the second-instance Board of Appeal are considered. These proceedings are also conducted in one of the three official languages of the EPO. In most cases, the vast majority of the actions filed before the EPO have the support of a qualified European patent attorney, who has analysed the application in order to generate the required documents in the language of proceedings with the necessary linguistic quality by the action.⁴⁰

Machine translation of patents is provided by the EPO’s Patent Translate system, and is useful not only to patent examiners, but also to its other users, especially those belonging to the patent-related legal setting. The service provided by Patent Translate was introduced for the purpose of making it possible for users to access

³⁶ Article 54 of the EPC.

³⁷ *Ibid.*

³⁸ Fromer (2009).

³⁹ Article 14 of the EPC. *See also* EPO, “Guidelines for Examination”, Chapter III, 14, available at: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/a_iii_14.htm (accessed 4 July 2018).

⁴⁰ Information collected directly at EPO’s official website: <https://www.epo.org/index.html> (accessed 4 July 2018).

the translation of patent documents into English (and the other official languages of the EPO) from as many as 31 other languages, which include Asian languages and Slavic languages, such as Russian.⁴¹ The access to these translations has changed the scenario of patent law, for example, by facilitating the processes of application and prosecution and reducing the costs of selecting documents through prior art search for further professional legal translation.⁴²

This growing importance becomes even more evident with the settlement of the European Patent with Unitary Effect. Since the Agreement on enhanced cooperation was signed, it demonstrated a concern with translation issues and included a very relevant article regarding the future of translation for the unitary patent. The document states that during the application process, demands for translations will be totally eliminated after a transitional period when machine translation tools will be sufficiently improved to assure efficient communication involving patent texts and the disclosure of the information they contain.⁴³

The following part of the document is worth transcribing:

The transitional period should terminate as soon as high-quality machine translations into all official languages of the Union are available, subject to a regular and objective evaluation of the quality by an independent expert committee established by the participating Member States in the framework of the European Patent Organisation and composed of the representatives of the EPO and the users of the European patent system. Given the state of technological development, the maximum period for the development of high quality machine translations cannot be considered to exceed 12 years. Consequently, the transitional period should lapse 12 years from the date of application of this Regulation, unless it has been decided to terminate that period earlier.⁴⁴

The translation regime of the Unitary System will considerably lower translation costs for applicants, as the national validation procedures will no longer be necessary, and therefore the charges associated with translation. It will also reduce the complexity related to national validation, as it offers protection covering up to 26 Member States of the European Union.⁴⁵ In the current system, the regimes of each country (including language and translation) can vary considerably, making it difficult for the applicant to apply for national validation in a larger number of countries. There are also cases of more than one language regime within the same country.⁴⁶ It is important to clarify that machine translation should not have legal

⁴¹ *Ibid.*

⁴² EPO, “A new benchmark for patent translation”, available at: <https://blog.epo.org/patent-information-2/new-benchmark-patent-translation/> (accessed 4 July 2018).

⁴³ See Sec. 2 of Council Regulation (EU) No. 1260/2012, *supra* note 5.

⁴⁴ *Ibid.*

⁴⁵ EPO, Unitary Patent, available at: <https://www.epo.org/law-practice/unitary/unitary-patent.html> (accessed 4 July 2018).

⁴⁶ In Belgium, for example, local companies should file the application in the predominant language of the region they are located, which is not the same for foreign companies. See “The world’s strangest patent filing requirements”, available at: <http://www.multiling.com/globalization-2/11-strangest-patent-filing-requirements-in-the-world> (accessed 4 July 2018).

effects in the new system, it being considered as a tool to facilitate the dissemination of patent information.⁴⁷

4 Patent Writing, Patent Translation and Disclosure

There is abundant literature on “how to write a patent”. Most of these works, which include books, manuals and articles on webpages, aim to help applicants increase the chances of having their patents granted and be “legally successful”, which means, to avoid further legal challenges. What these works probably have in common is the fact that they present the task of writing a patent as a very complex one, which not only requires technical knowledge and a capacity to explain the content of an invention, but demands adapting the text to specific language rules.⁴⁸

In this regard, these works clearly show that the success in having a patent granted and avoiding opposition and disputes depends, to a certain extent, on effective writing. However, defining what effective writing is for patent documents is also more complicated than it seems at first sight, and may depend directly on which national system the application is directed to, as well as which type of technology and scope of protection, among the other mentioned factors.

Sheldon highlights the complexity writing a patent application stating that there are few tasks more difficult than that of preparing a patent application for an invention. Curiously, the same author describes writing a patent as an art, comparing it to the routine of learning how to play a musical instrument. According to Sheldon, only talented attorneys are able to write successful patents.⁴⁹

According to WIPO’s “Patent Drafting Manual”, a patent application is divided into claims, detailed description (or specification), drawings, background, abstract and summary. According to the Manual, the claims are the first part to be written, and should clearly show that the agent “understood” the invention.⁵⁰ For DeMatteis et al., the claims are the most crucial part of a patent with respect to its grant and “legal success”. They explain that the inventor should write the very first draft of what will be the patent in simple language, having the challenge to describe the invention in a way that makes it easier for the practitioner to “translate” it into patent legal language.⁵¹

The “Guide for applicants”, published by the EPO, also presents instructions on the writing of a patent.⁵² Articles 83 to 85 and Rules 42, 43, 47 and 48 of the EPC detail the requirements associated with the content of the description, claims,

⁴⁷ Council Regulation (EU) No. 1260/2012 (12), *supra* note 5.

⁴⁸ See the following examples: Sheldon (2015); DeMatteis et al. (2006); Pressman (1985); Waller (2011). USPTO. “Design patent application guide”. Available at: <https://www.uspto.gov/patents-getting-started/patent-basics/types-patent-applications/design-patent-application-guide> (accessed 4 July 2018).

⁴⁹ Sheldon (2015), pp. 1-1–1-2. See *supra* note 48.

⁵⁰ “WIPO Patent Drafting Manual”, p. 33. Available at: http://www.wipo.int/edocs/pubdocs/en/patents/867/wipo_pub_867.pdf (accessed 4 July 2018).

⁵¹ DeMatteis et al. (2006), *supra* note 48, p.183.

⁵² EPO, “Guide for applicants: how to get a European patent”. Available at: https://www.epo.org/applying/european/Guide-for-applicants/html/e/ga_c_ii_3.html (accessed 4 July 2018).

drawings and abstract. According to Art. 83, the requirement of disclosure of the invention is a central one to the patent text. Article 85 determines that the abstract should “serve the purpose of technical information only”, and may not be considered for any other purpose, such as for interpreting the scope of protection. The claims, on the other hand, should “define the matter for which protection is sought”.⁵³ In addition, they should “be clear and concise and be supported by the description”.⁵⁴

Rule 43 of the EPC regulates the form and content of claims. It shows that the claims are deeply related to the matter of the invention and to the scope of protection. Legally, the claims are the most important part of a patent and for the disclosure of its content to patent examiners and legal players. According to Rule 43, claims shall comply with a rigid text structure, which encompasses:

[...] a statement indicating the designation of the subject-matter of the invention and those technical features which are necessary for the definition of the claimed subject-matter but which, in combination, form part of the prior art; [...] a characterising portion, beginning with the expression “characterised in that” or “characterised by” and specifying the technical features for which, in combination with the features stated under sub-paragraph (a), protection is sought.⁵⁵

No doubt writing accurate patent applications is a task for skilled, experienced practitioners, who may be represented by different professionals or firms directed to distinct technologies, and even to distinct levels of experience within the targeted national or regional system. In this regard, Sheldon also warns of the risk of additional costs and the risk of further legal problems. For Sheldon, whose work comprises an important reference on patent writing, the application, even when filed only at the national level, should be written with “an eye overseas”, as it may later be a vehicle for obtaining foreign patents.⁵⁶

In this context, each foreign application will follow the national or regional office’s language and translation regime. In the case of the European patent, the application must be filed in one of the EPO’s official languages. If the application language is not represented by one of these official languages, it should be translated into one of them. The language of the application or the language of translation, should one be necessary, will be considered as the language of the proceedings in all further proceedings before the EPO.⁵⁷

These regimes and their detailed regulations concerning language and translation allow the assertion that patents should be written to be translated, even when they are first filed only nationally. The possibility of foreign applications is something that is recommended for inventors to take into account when preparing the first

⁵³ EPC, Art. 83 to 85 and Rules 42, 43, 47 and 48.

⁵⁴ EPC, Art. 84.

⁵⁵ Rule 43 of the EPC. Available at: <https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/r43.html> (accessed 4 July 2018).

⁵⁶ Sheldon (2015), *supra* note 48, pp. 1–5.

⁵⁷ EPC, Art. 14(c).

application. This requires an even more skilled practitioner or attorney, who should describe the idea of the invention in legal patent language. Theoretically, this agent should work with clear, translatable texts.⁵⁸

Clear writing remains important when the patent is not going to be immediately or necessarily translated into other languages. In this regard, it is worth acknowledging that patents are published in multilingual platforms. The patent system has a global reach, and patents are mainly accessible to non-native speakers of the languages in which they were published through machine translation. The more “translatable” the text is, the more it will also disclose the patent content through its machine-translated versions. Therefore, as machine translation supports prior art search and examination, clear writing and translation may have legal implications.⁵⁹

Another reason why patents should be written with the possibility of further translation in mind is the fact that translations represent a major expense in the filing and prosecution of foreign patents, and using certain strategies in writing can reduce these costs. According to Sheldon, there should be more incentive in avoiding verbosity, repetition and circumlocution in a foreign application than in a national one. Further, because this can decrease the costs of a translation considerably, it would also save time and money if those deficits were avoided in the national original version of a patent.⁶⁰

Sheldon then points to clarity, consistency and brevity as very important features to bear in mind while writing a patent. These features are also indicators of translatability.⁶¹ The more clear, accurate and concise the text is, the better the chances are of obtaining appropriate translations from it. He then advises using clear terms and explanations when the meaning of a term is unavoidably vague. He also suggests using certain techniques to elaborate a well-structured, “organized” text, such as numbered or lettered paragraphs and simple syntax. Sheldon finally suggests the strategy of reverse translation to prove the level of “clarity” of the patent text. This means the text should be translated into a target language (for future application, for example) and then retranslated into English.⁶²

The instructions of the abovementioned literature on how to write a patent application are valid under the idea that a patent text is meant to be searched and accessed universally in patent databases. Besides, patent prior art search constitutes the basis of patent examination and may ground decision-making during the entire process of prosecution and in legal disputes. It is also crucial with respect to the duty of the applicant to provide information on prior art.⁶³

For these reasons, patents should be available in clear language, in a way that machine translation will suffice in disclosing its contents to the person skilled in the

⁵⁸ Sheldon (2015), *supra* note 48, pp. 18-03–18-22.

⁵⁹ The EPO’s “Guidelines for Examination” acknowledge the legal relevance of machine translation and its wide use for prior art search. For the complete reference, *see supra* note 6.

⁶⁰ Sheldon (2015), *supra* note 48, pp. 18-10–18-11.

⁶¹ For an overview of the concept of “translatability”, *see de Pedro* (1999).

⁶² *Ibid.*

⁶³ Articles 90 and 124 of the EPC.

art in order to assure legal certainty within the patent system. For the applicant, writing a clear text also represents a way of avoiding further unnecessary legal challenges that may be related to inaccurate translations.⁶⁴

5 Machine Translation Discloses Patent Information

Inventions are associated with both scientific and economic development. They represent a specific type of intellectual creation, which is protected mainly by patent rights. For this reason, the patent document, which is related to an invention, should consistently describe its subject matter in order that this content can be easily accessed in its original language and in translation, when searched for in patent databases. The patent text should have these features because of the condition of disclosure of technological information, which represents the main economic justification to give temporary monopolies to inventors.⁶⁵ That is why the application should disclose the invention in a clear and complete manner.⁶⁶

Patent law associates disclosure with the criteria of patentability required by the grant proceedings of patent examination. The EPO's "Guidelines for Examination" describe the disclosure of the invention as a "further requirement" of patentability, together with the basic requirements of novelty, industrial application and inventive step, "The invention must be such that it can be carried out by a person skilled in the art (after proper instruction by the application)".⁶⁷

As discussed in Sect. 4 of this article, adequate writing impacts the disclosure of the content of a patent, and also the translation of its text. In addition, in the current association between writing, disclosure and translation, machine translation plays a very important role, representing the primary basis of communication to the global patent system. The role played by machine translation has increased as its quality has improved and the volume of patent documents available in search platforms in different languages has expanded.

The results of the present research intend to contribute to a better understanding of the role played by machine translation in the current patent system in terms of the level of disclosure it really makes possible. The first and most relevant conclusion, made possible by the analysis of the sample of patents, is an affirmative answer to the research question, providing that machine translation clearly discloses the contents of inventions.

The average score, referring to number of errors, for the first set of languages (excluding Chinese), was 22.75, which illustrates that the content is appropriately disclosed through machine translation in an average of 77.25. In other words, almost

⁶⁴ Sheldon (2015), p. 18-03, *supra* note 48.

⁶⁵ On the relationship between disclosure and the justification of the patent system, *see* Landes and Posner *supra* note 26, pp. 294–327.

⁶⁶ EPC, Art. 83.

⁶⁷ EPO, "Guidelines for Examination", Chapter I – Patentability. 1 – Basic Requirements and 2 – Further Requirements. Available at: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_i_2.htm (accessed 4 July 2018).

80% of the content is disclosed. For the Chinese patents, the average was 68%, which is lower than the results for the other languages, as expected, but still presents a reasonable level of quality.⁶⁸

The results also make clear that the information is disclosed under the following circumstances: (1) even if the languages involved in the translated text are distant in terms of linguistic origin (e.g. Chinese and Portuguese); (2) if the person skilled in the art does not have any knowledge of the original language; (3) were the texts long or short and without any aid of human translation; and (4) if the person skilled in the art accesses only the texts of the patents, without following the drawings and the description that follow the claims, which would make the understanding much easier.

However, there are still some limitations on patent disclosure through machine translation. One of them is the fact that databases or patent offices' searching tools offer different possibilities, sometimes divergent from each other, of translation machines. This continues to make machine translation services related to searching tools limited and susceptible to improvements.⁶⁹

Searching for a patent in a foreign language may still be a difficult task, in spite of the significant improvements observed during the last decade. Linguistic proximity is an important factor, because technical terms tend to be similar in languages belonging to the same family. That became clear through the analysis of the patents in French, Italian and Spanish, having Portuguese as the target language. Although the translation may present the same average score of other (non-Latin) languages, the similarity still favours understanding.

One reason for some poor translations involving close languages is the fact that English still intermediates patent machine translation, including the ones related to languages as close as Spanish and Portuguese, working as a hub in statistical machine translation. This is an important factor which affects the value of the translated text in disclosing the contents of the patent. As a result of the analysis, there were frequent cases in which cognates or sentence structures that could be easily translated from Spanish, Italian or French, into Portuguese, were inconsistently translated, as an effect of the indirect translation.⁷⁰ In this regard, it is important to add that this reality is changing fast, with the advent of neural technologies for machine translation, which are capable of combining the memory information of a broader network of languages.⁷¹

Other aspects that should be considered, and can be subject to further (more detailed) studies, are the structure and quality of the original texts. The French patents obtained the best score according to the LISA-QA method, reaching an average of only 14.25 errors, which can be considered as a translation of very high quality. This shows that the level of quality of a translation depends not only on linguistic similarity, although this is also a factor to be considered. In the case of

⁶⁸ It is important to emphasise that the analysis was not limited to translation quality assessment, but included a second step represented by a survey filed by collaborators representing the person skilled in the art, as explained in Sect. 2.

⁶⁹ Further studies are recommended for comparing the results of different translation machines.

⁷⁰ Brown et al. (1990), pp. 79–85.

⁷¹ Koehn (2017)

Figure 3 Average score obtained by each language (or group of language) according to LISA-QA. Source: based on the results of LISA-QA's analysis.

French patents, the clearness of the original texts may have played an important role, their writing being characterised by short and simple sentences and adequate text organisation (parallelism to facilitate understanding, accuracy in the usage of technical terminology, general coherence among the sentences, etc.). The graph above shows the average score in number of errors obtained by each language or group of languages (Fig. 3).

German was the language with the highest results in number of errors (26.25), among the Western languages, which may be due to the factor of linguistic distance from the target language, but also to the fact that some patents displayed long sentences (in long texts), complex sentence structure with numerous subordinate clauses and the intricate use of adjectives. Still the difference between the lowest and highest results (in this case, French and German patents) cannot be considered as a large variation, being represented by 12 points only. The other scores were 22.75 for English patents, 19.5 for Spanish patents, and 24.00 for Italian patents.

In the case of both French and German patents, all of the inventors clearly had French and German national origins, which was not the case with the English patents, whose number of foreign applicants was clearly considerable. Although the analysis did not focus on the relation between the nationality of the applicant and the language of the patent, this could be the subject of further studies.

There was not a considerable difference in the scores obtained by types of technologies. The results were similar to solar and wind power technologies. The results were also similar for abstracts and claims, with a small advantage for the claims in terms of quality of the translation (the average score for the abstracts was 26.5 and the average score for the claims was 22.75). This can be explained by the fact that claims tend to be more technical than abstracts, presenting the inventions in a more detailed and structured way. The fact that abstracts are to present the contents of the patents in a concise way may result in a less clear text. In any case, the results are positive in terms of the level of disclosure, as the claims are the essential part of a patent for disclosing its content.⁷²

⁷² On the function of the claims for disclosure of patent information, *see supra* note 23.

By using the method based on the LISA-QA score table, it was possible to identify the types of errors that interfere most seriously in the quality of patent machine translation. The types of errors analysed through this methodology are: mistranslation, accuracy, terminology, language (syntax) and style. These errors were categorised (and weighted by 1, 5 or 10, respectively) as minor, major and critical. The results showed that the variety of errors tends to diminish as the degree of severity of the error grows. That may also have positive indications to machine translation as a means to disclose patent information, as the largest varieties of errors do not occur in the critical level of severity. The results also showed that language structure is the most important challenge in improving machine translation, as all the languages presented a considerable number of language (syntax) errors on all levels, notably in German patents.

As explained in Sect. 2, the methodology used presents a certain degree of subjectivity, as it is not different, in this sense, from any other method for translation quality assessment. However, if patent machine translation is used as a tool by a person skilled in the art, whose aim is to clearly understand the content of the invention, its results are applicable and may actually suffice for the purpose of bringing new light into the debate over machine translation and disclosure of patent information in patent law.

The survey with the collaborators, representing the persons skilled in the art,⁷³ confirmed the results of LISA-QA. As explained in Sect. 2, the collaborators were given three patents (abstract and claims) machine translated into Portuguese from three languages (English, German and Chinese), without knowing the original language. They were unanimous in concluding that the content of the three inventions was disclosed through the text they read. Their opinions were similar about the three texts.

In addition, the collaborators commented that, although the texts were confusing in some passages, it was still possible to understand the technical content. They also considered the text translated from German to be the most difficult one to understand, due to the reference to drawings that were not supplied. They added that the understanding would have been much easier with the drawings. They also commented that the text translated from English was the clearest one and that the text translated from Chinese could be understood after a certain effort.

6 Conclusions and Recommendations for Further Studies

The main conclusion of the present work is that machine translation discloses patent information. Although this conclusion calls for further studies, it still represents an important starting point to contextualise and reframe the importance of machine translation to the patent system. It is believed that the perception of the role of machine translation in the disclosure of patent information is still frequently based on opinions and prejudices, which can harm a rational interpretation of its use when planning new cross-boarder agreements or policies.

⁷³ For the concept of “the person skilled in the art”, *see supra* note 8.

This research has also shown the level of disclosure that machine translation is capable of reaching: almost 80% for Western languages and almost 70% for Chinese-English. The greatest challenge to improvement is still in language structure (syntax). A second challenge involves the translation of technical terms. However, these challenges do not diminish the significance of the quality of machine translation reached until now. Although machine translation can still improve, it clearly discloses patent information and represents one of the main bases of communication within the patent system.

The results confirm the idea that languages no longer represent an insurmountable barrier to the disclosure of patent information. Users of the patent system can rely on machine translation in searches for prior art and for a basic understanding of the patent text, which may be supplemented with professional human translation when necessary. In this regard, cross-border agreements, such as the enhanced cooperation in the area of the creation of a unitary patent in Europe are following an updated view of the patent system when relying on machine translation as the main means to disclose patent information.

There are various possibilities of future studies concerning (human and machine) translation and the patent system. The most relevant of them can be summed up in the following questions: What are the effects of regional agreements to harmonise language regulations and reduce costs with translation? Can companies that provide language and translation services make errors (for example, during previous searches for documents through machine translation) or produce translations that could bias the legal path of a patent? Does translation influence the results of patent examination? What is the role of the patent system in the evolution of machine translation?

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Berg B (2001) Qualitative research methods for the social sciences, 4th edn. Allyn & Bacon, Boston, p 255
- Brown et al (1990) A statistical approach to machine translation. *Computational linguistics* 16(2):79–85
- Cyril G et al (2010) Learning machine translation. MIT Press, Cambridge
- DeMatteis B, Gibbs A, Neustel M (2006) The patent writer: how to write successful patent applications. Square One Publishers, New Hyde Park
- Gibbs and DeMatteis (2003) Essentials of Patents. PatentCafe, p 14
- Fromer (2009) Patent disclosure, 94 *IOWA L. Rev.* 539, 606, pp. 544–560, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1116020. Accessed 4 July 2018
- Koehn P (2017) Statistical machine translation draft of chapter 13 – neural machine translation. Johns Hopkins University, p 6. Available at: <https://arxiv.org/pdf/1709.07809.pdf>. Accessed 4 July 2018
- Krippendorff K (2004) Content analysis: an introduction to its methodology, 2nd edn. Sage, Thousand Oaks
- Landes W, Posner R (2003) The economic structure of intellectual property law. Harvard University Press, Cambridge

- Nam Y, Barnett G (2011) Globalization of technology: network analysis of global patents and trademarks. *Technological Forecasting and Social Change* 78(8):1471–1485
- de Pedro R (1999) The translatability of texts: a historical overview. *Érudit* 44(4)
- Pressman D (1985) *Patent it yourself*. Nolo Press, Berkeley
- Sheldon J (2015) *How to write a patent application*, 3rd edn. Practising Law Institute, New York
- Waller F (2011) *Writing chemistry patents and intellectual property: a practical guide*. Wiley, Hoboken
- White J (2003) How to evaluate machine translation. In: Somers H (ed) *Computers and translation: a translator's guide*. John Benjamins, Amsterdam, pp 211–244