

## **Generating a classification for EUIPO trademark filings – A string matching approach**

Peter Neuhäusler<sup>1,2</sup>, Alexander Feidenheimer<sup>1</sup>, Rainer Frietsch<sup>1</sup>, Henning Kroll<sup>1,3</sup>

<sup>1</sup> Fraunhofer Institute for Systems and Innovation Research ISI, Competence Center Policy – Industry – Innovation, Breslauer Strasse 48, 76139 Karlsruhe, Germany

<sup>2</sup> Berlin University of Technology, Chair of Innovation Economics, VWS 2, Müller-Breslau-Strasse, 10623 Berlin, Germany

<sup>3</sup> Institute of Economic and Cultural Geography, Leibniz Universität Hannover (LUH), Schneiderberg 50, 30167 Hannover, Germany

## Contents

<b>Abstract</b> .....	<b>1</b>
<b>1 Introduction</b> .....	<b>1</b>
<b>2 Existing trademark classifications - a brief overview of the literature</b> .....	<b>3</b>
<b>3 The approach</b> .....	<b>6</b>
3.1 From textual descriptions to a classification.....	6
3.2 Data preparation & matching .....	10
<b>4 Results</b> .....	<b>11</b>
4.1 Coverage and validation of the matching .....	11
4.2 The classification setup .....	15
4.3 Examples and use cases .....	16
4.4 Pitfalls.....	20
<b>5 Conclusions</b> .....	<b>21</b>
<b>6 Appendix</b> .....	<b>22</b>
<b>7 References</b> .....	<b>37</b>

## Tables

Table 1:	Definition of technology related NICE-classes regarding goods .....	5
Table 2:	Exemplary trademark descriptions .....	7
Table 3:	Classification scheme for Class 12 "Vehicles; apparatus for locomotion by land, air or water" .....	9
Table 4:	Lemmatized descriptions of trademarks .....	10
Table 5:	The trademark classification at level I.....	23

## Figures

Figure 1:	Explanatory Note to the NICE classification - excerpt for class 12 ("Vehicles; apparatus for locomotion by land, air or water").....	8
Figure 2:	Share of matched trademarks .....	12
Figure 3:	Share of matched trademarks by country and NICE class.....	13
Figure 4:	The classification setup.....	16
Figure 5:	Trademark classification at Level I, World, 2018, absolute numbers.....	17
Figure 6:	Shares of trademarks in total trademarks at classification level I, China, Germany and the U.S., 2018 .....	18
Figure 7:	Shares of trademarks within level I class "information technology and audio-visual, multimedia and photographic devices" at level V, world, 2018 .....	19
Figure 8:	Shares of trademarks within level I class "information technology and audio-visual, multimedia and photographic devices" at level V, Germany, 2018 .....	20

## Abstract

This paper aims to analyze topics within the international classification of goods and services (NICE classes) applied for the registration of trademarks at the EUIPO. This is accomplished by introducing a more fine-grained classification of trademarks as a "subsection" of the rather rough NICE classes. To do this, we relate the descriptions of the trademarks that the applicants provide upon filing to the list of pre-defined keywords that are available from the WIPO to assist the applicant in describing his or her mark. In order to relate the keywords to the classifications, i.e. to assign trademarks to the classification, we use two algorithms including a Levenshtein-based matching and a Jaro-Winkler algorithm based matching. The Levenshtein-based approach already leads to a coverage of 75% of matched trademarks. With the help of the Jaro-Winkler matching algorithm (in combination with the Levenshtein distance) we could assign another 10%, leading to a coverage of 85% of all EUIPO trademarks matched to at least one classification key in 2018. Based on this matching we generate a hierarchical classification including five layers, the first layer including 234 classes up to the 5<sup>th</sup> layer which comprises 8,613 distinct classes.

## 1 Introduction

In quantitative innovation research, trademarks are more and more extensively used for the measurement of innovation activities (deGrazia et al., 2020). This is particularly true at the micro level, where the relationship between trademarks and innovation has been well established (e.g. Flikkema et al., 2014; Greenhalgh and Rogers, 2006; Sandner and Block, 2011). Some more recent studies, however, have also focused on the macro level, especially on the relation between trademarks and trade, e.g. by analyzing the effect of the strength of trademark protection in a country on exports of domestic firms (Yang et al., 2018) or whether trademarks by foreign firms are connected with high-tech imports (Perepechko, 2018). DeGrazia et al. have recently also provided evidence that aggregate trademark data could predict business cycles (deGrazia et al., 2020). In sum, the body of literature on trademarks as innovation indicators has grown, especially in the recent ten years.

This has several reasons. Although trademarks can be filed for products like technical equipment or technical procedures, also services are eligible for protection within the system of trademark rights. Consequently, trademarks provide additional insights on service sector innovation and can thus serve as a complementary indicator to patents, which are focused on technical inventions (Gauch, 2007; Sandner and Block, 2011; Schmoch, 2014). Trademarks have been shown to be well applicable as a measure of innovative

outcomes in the case of knowledge-intensive business services trademarks (Schmoch and Gauch, 2009) and they gain additional relevance through the fact also a registration of copyrights is not possible in many jurisdictions (Schmoch, 2003). Trademarks are also often used as a valuable complement within a firm's IPR portfolio (Sandner and Block, 2011).

Furthermore, data availability has improved notably in recent years and data access was massively eased. The publication of large scale datasets, for instance by the European Union Intellectual Property Office (EUIPO) or the United States Patent and Trademark Office (USPTO) has opened new opportunities to access and use trademark data (Castaldi, 2019). Finally, also the technical possibilities of analyzing textual data have increased and with it the ease of analyzing unstructured data by non-experts, which enables analyses on trademarks that have not been possible before.

The trademark indicator has several advantages. Often, trademarks protect names of products (and not whole firms) and thus can be seen as an innovation indicator that is rather close to the market (Mendonca et al., 2004; Schmoch, 2014). Furthermore, trademarks are a very timely indicator of innovation activities (Thoma, 2019), in the sense of being available much more quickly than patent indicators after filing, which makes them viable as an early-stage innovation indicator (although trademarks are less close to the invention in the innovation process itself). In combination with other IPR related indicators, e.g. patents, industry designs or utility models, they further allow us to look at effects of IPR portfolios and their combined effects on innovation.

However, there also are downsides to trademarks as an indicator of innovative activities. What limits the value of trademark data in innovation studies is that there are yet no guidelines to ascertain which trademarks relate to innovation. Although there are estimates that 60% of registered trademarks refer to innovation activity (Flikkema et al., 2014), the identification of innovation-related trademarks still remains a problem (Flikkema et al., 2019).<sup>1</sup> This stems from the fact that trademarks are an "untested right", i.e. they are not content certified and checked for the criteria of "inventive step" like patents are. Only formal criteria are checked upon filing, meaning that trademarks are only registered but not examined. Furthermore, the costs for filing a trademark are comparably limited, so the entry barriers for applicants to file a trademark are rather low.

A major step towards solving this problem is to try to find out more about the content of a trademark. This is what we aim for in this paper by creating an in-depth trademark

---

<sup>1</sup> The study of Flikkema et al. (2019) referred to small- and medium-sized enterprises registered in Benelux countries.

classification that goes beyond the current international classification of goods and services in trademarks (NICE classification) in terms of detail. We make use of the trademark description in trademarks filed at European Union Intellectual Property Office (EUIPO) that each applicant has to provide when applying for the EU trademark. Since this is a semi-standardized process - the applicant chooses from a given list of keywords per trademark class but may also choose his own words - it allows us to set up such a classification. First of all, this allows us to perform more detailed, automated analyses of specific fields where trademarks are filed, e.g. IT services, electric mobility etc. In addition, it can serve as a basic step towards solving the problem of the identification of innovation-related trademarks, as more detailed analyses are possible once information on the content of a trademark is more easily accessible.

In the following sections, we will provide a brief overview on the literature of how trademarks are classified and which earlier papers have tried to deal with the problem of identifying innovation-related trademarks. Second, we will provide a detailed overview of the methodology for the creation of the classification. Then, we will describe the results and give an overview on what can be achieved with the new classification.

## **2 Existing trademark classifications - a brief overview of the literature**

Among formal instruments of intellectual property protection, trademarks are in widespread use. All "tokens", e.g. words, pictures, that are suitable to distinguish a company's goods or services from those of other companies are eligible for protection. These can for example be words, individual letters, numbers, pictures, colors and even acoustic signals. Trademarks are valid for ten years after filing and can be renewed indefinitely (Deutsches Patent- und Markenamt (DPMA) 2008; Graham and Harhoff 2006).

There are various ways to obtain trademark protection. In order to achieve protection in a certain country, the national patent and trademark office (for instance the German Patent and Trademark Office (DPMA)) is possible. Alternatively, the registration of a EU Trademark, which is valid across the whole EU, or the registration of an international trademark at the WIPO, which is valid in all countries who have signed the Madrid Protocol, are possible options.

Upon receipt of an application of a trademark at a given trademark office, the trademark will be processed. This step includes a classification of the trademark according to several classifications, a check of formalities, and a check of the trademark "on absolute grounds", i.e. the trademark is analyzed to see whether it is distinctive but not descriptive, translation as well as a search for identical or similar marks including a "surveillance

letter" that informs third parties about the filing of the given trademark (Office for the Harmonization of the Internal Market (OHIM) 2014). As already mentioned in the introduction, this means that, different to patents, trademarks are not content certified. Only formal criteria are checked upon filing. The pursuit of potential violations or infringements of registered trademark rights lies in the hands of the trademark owner. A procedure of cancelling a competing trademark can be initiated only if a trademark holder indicates a violation. After the examination period, a trademark is published. From the date of publication, third parties have three months to object to the registration of the trademark either based on "earlier rights" or on "absolute grounds". If nobody files an opposition, the trademark is registered and the registration is published. After registration, only official appeals can be used to challenge the official decision by the OHIM (Office for the Harmonization of the Internal Market (OHIM) 2014).

A trademark is only valid or effective in the pre-defined business circles or business areas. These areas are defined by the classes of goods and services and the further specifications within the classes. Based on this, distinctiveness is achieved. In consequence, a word mark like "BOSS" can be used exclusively for clothing and might also distinctively mark a text-marker by another company as the business cycles are distinct and a distinctiveness according to the use of the products is given. In this respect, the classes and the descriptions are similar to the claims of a patent, which define the subject and range of the patent or the trademark, respectively. If applicants aim to achieve broad coverage in a large number of business areas, then they need to register in a large number of classes. For example, a soft-drink producer might not only want to achieve distinctiveness in the business area of soft-drinks, but also sell crockery, clothing, household appliances etc. with the same logo and also achieve distinctiveness. The applicant would then need to register the particular trademark in all the respective classes.

Upon application, most trademark offices allow for the registration in one class and offer up to three classes at no or low additional costs, while the fees increase with the number of registered classes. The question of a broad coverage is therefore also a question of financial resources or - to be more precise - of cost-benefit ratios. Furthermore, the broader the coverage, both in terms of classes and in terms of the listed goods or services, the higher is the probability of the existence of a confusing similarity and therefore an opposition to the particular trademark application by third parties. It is thus costly and risky to name too many classes and goods or service categories. In addition, after five years, the applicant needs to prove the use of the trademark in general as well as in the particular classes/business areas. If the applicant fails to do so, the trademark might be deleted or at least restricted. These are the reasons why it is reasonable to infer the content/subject of the trademark filing from its classification.



For trademarks, several national and international classifications exist. The two systems applied by the EUIPO are the Vienna Classification (VCL) and the NICE Classification (NICE). The Vienna Classification (VCL) is an international classification of the figurative elements of marks. It was established by the Vienna Agreement in 1973 and is currently available in its 8th version (WIPO <https://www.wipo.int/classifications/vienna/en/>). More important for our purposes, however, is the NICE Classification also known as the International Classification of Goods and Services. It was established by the Nice Agreement in 1957 and is open to states who are parties to the "Paris Convention for the Protection of Industrial Property" that already originated in 1883 (currently including 84 countries as well as the "World Intellectual Property Organization (WIPO), who specifies the NICE classes) (WIPO <https://www.wipo.int/treaties/en/classification/nice/>). The NICE Classification is updated every five years - with minor annual changes - and is currently in its 11th version. It consists of 45 classes, of which classes 1 to 34 are assigned to goods and classes 35 to 45 are assigned to services (EUIPO <https://euipo.europa.eu/ohimportal/en/nice-classification>). The classes define the scope and the context of each application and are provided by the applicants themselves.

Due to the differentiation in product marks and service marks within the classification, it is relatively easy to analyze product marks, service marks and mixed product/service marks, i.e. marks that are assigned NICE classes referring to goods as well as NICE classes referring to services, independently. However, this does not lead to a differentiation between innovative and non-innovative trademarks. There have been several approaches in the literature to achieve this aim. Earlier studies by Schmoch and Gauch. (Gauch, 2007; Schmoch, 2003; Schmoch and Gauch, 2009) applied a differentiation of "research-intensive services" with regard to service marks by where the classes 35, 36, 38, 41, 42, 43, 44, 45 were defined as research-intensive services. In the case of products, they concentrated on eight fields that have been defined as having high technology relatedness. These are summarized in Table 1.

Table 1: Definition of technology related NICE-classes regarding goods

No.	Name	NICE classes
1	Chemistry	1, 2, 3, 4, 13
2	Pharmaceuticals	5
3	Metals	6
4	Machines	7, 8
5	Electronics (components, instruments)	9, 14
6	Medical technologies	10
7	Electronic devices	11
8	Vehicles	12

Source: Schmoch (2003)

### 3 The approach

The basic approach for the generation of our classification of EUIPO trademarks follows a relatively simple idea. Upon filing a trademark at the EUIPO, an applicant needs to characterize it by providing a keyword based trademark description for each chosen trademark class. To assist applicants in doing this, the WIPO publishes "Explanatory Notes" to the NICE classification, from which the applicant can choose respective terms for his description (WIPO, 2018). To further aid the applicant, he or she can also use TMClass, an online classification assistance of the EUIPO (<https://euipo.europa.eu/ec2/>).<sup>2</sup> Besides these options, the applicant can also use self-selected keywords to describe his trademark.

These descriptive keywords, however, are not related to the explanatory notes in the NICE classification. We fill this gap by providing a string-matching algorithm based on the Levenshtein (Navarro, 2001) and Jaro-Winkler distance (Winkler, 1990) to relate the description provided by the applicant to the standardized keywords stated in the explanatory note to the NICE classification. The explanatory note already includes a hierarchical set-up in the form of a basic taxonomy, with sections and sub-sections, which avoids one of the major problems in the generation of taxonomies (Krishnapuram and Kumnamuru, 2003).

#### 3.1 From textual descriptions to a classification

A description of the content of the trademark, like an abstract or an in-depth description of an invention, as in the case of patents, is not available for trademarks. Keyword descriptions are the only text that is available for their characterization. As already stated above, an applicant has to assign one or several NICE classes to his applications upon filing. For each NICE class, it is necessary to provide one or several keywords to describe the trademark in more detail. For this purpose, applicants can use a pre-defined list of keywords that they can assign to their trademark, which is the method of choice in the majority of cases. However, they are also free to use their own words to describe the trademark. An example of the textual description of EUIPO trademarks is provided in Table 2.

---

<sup>2</sup> In TMClass, the applicant searches for a term and gets the corresponding Nice class as well as related terms for the filed trademark. The terms can also be translated to other languages.

Table 2: Exemplary trademark descriptions

Application no.	Class no.	Goods/Services description
17836735	12	Vehicles; Apparatus for locomotion by land; Cars; Parts and fittings for vehicles.
17833096	12	Motorized vehicles, other than rail vehicles; Motors, engines and drives for motor vehicles; Propulsion mechanisms for automobiles; Chassis for motor vehicles.
17814559	12	Vehicles and conveyances.
17812801	12	Bicycles; freewheels for bicycles; bicycle wheel hubs; bicycle chains; bicycle gears; derailleurs for bicycles; Transmission belts for land vehicles; power transmission belts for land vehicles; tires; anti-theft devices for vehicles.

Source: EUIPO.

In order to help the applicant to define his trademark, the WIPO publishes the "Explanatory Notes" to the NICE classification, which can be used by the applicant to select pre-defined keywords. These explanatory notes are set-up hierarchically and can therefore be used to generate a hierarchical classification for trademarks.

In Figure 1, an excerpt of the explanatory notes to the NICE classification for class 12 "Vehicles; apparatus for locomotion by land, air or water" is provided. As can be seen from the figure, class 12 has sections ("vehicles and conveyances"), which have further sub-categories that can be used as sub-sections ("land vehicles and conveyances"). At the lowest layer, there are terms for the description of the trademark that can be applied by the applicant to characterize his trademark, which already have a unique identification number.

Figure 1: Explanatory Note to the NICE classification - excerpt for class 12 ("Vehicles; apparatus for locomotion by land, air or water")

Basic No.	Indication
	• <b>VEHICLES AND CONVEYANCES</b>
120004	<b>aerial</b> conveyors
120237	<b>air</b> cushion vehicles
120188	<b>cable</b> transport apparatus and installations
120051	<b>cars</b> for cable transport installations
120180	<b>chairlifts</b>
120144	<b>military</b> vehicles for transport
120170	<b>ski</b> lifts
120110	electric <b>vehicles</b>
120193	<b>vehicles</b> for locomotion by land, air, water or rail
120149	<b>water</b> vehicles
	•• <b>LAND VEHICLES AND CONVEYANCES</b>
120233	<b>ambulances</b>
120271	<b>armoured</b> vehicles / <b>armored</b> vehicles
120310	<b>bicycle</b> trailers
120044	<b>bicycles</b>

Source: WIPO, 2018

Based on this explanatory note, we have generated a classification scheme with five layers (Level I to Level V), with the 5th layer being the respective term from the trademark description and a description identifier (L5-Id). An example from the classification is shown in Table 3 for NICE Class 12 "Vehicles; apparatus for locomotion by land, air or water". Basically, we have translated the Explanatory Note to the NICE classification to a machine readable format. Special characters were excluded. Spelling variations were stored in multiple lines using the same class identification number.

Table 3: Classification scheme for Class 12 "Vehicles; apparatus for locomotion by land, air or water"

No.	Layer I	Layer II	Layer III	Layer IV	Layer V	L5-Id
12	VEHICLES AND CONVEYANCES	LAND VEHICLES AND CONVEYANCES	GENERAL	GENERAL	cars / automobiles	120199
12	VEHICLES AND CONVEYANCES	LAND VEHICLES AND CONVEYANCES	GENERAL	GENERAL	camping cars / motor homes	120249
12	VEHICLES AND CONVEYANCES	LAND VEHICLES AND CONVEYANCES	GENERAL	GENERAL	driverless cars [autonomous cars]	120279
12	VEHICLES AND CONVEYANCES	PARTS AND FITTINGS FOR VEHICLES	GENERAL	GENERAL	vehicle chassis	120069
12	VEHICLES AND CONVEYANCES	LAND VEHICLES AND CONVEYANCES	GENERAL	GENERAL	bicycles	120044
12	VEHICLES AND CONVEYANCES	PARTS AND FITTINGS FOR VEHICLES	WHEELS AND TYRES	GENERAL	freewheels for land vehicles	120148
12	VEHICLES AND CONVEYANCES	PARTS AND FITTINGS FOR VEHICLES	ANTI-THEFT, SECURITY	GENERAL	safety belts for vehicle seats	120059
12	VEHICLES AND CONVEYANCES	PARTS AND FITTINGS FOR VEHICLES	ANTI-THEFT, SECURITY	GENERAL	anti-theft devices for vehicles	120200
12	VEHICLES AND CONVEYANCES	PARTS AND FITTINGS FOR VEHICLES	PARTS FOR LAND VEHICLES	GENERAL	bicycle chains	120061

Source: WIPO, 2018, own attributions

The challenge now is to relate the trademark descriptions to the classification scheme, i.e. to combine Table 2 and Table 3 to assign the keywords in the trademark descriptions to the classification scheme so that each trademark is assigned to one or more classes in the classification.

### 3.2 Data preparation & matching

Before we can start with the matching of the keywords in the classification to the keywords in the description, we have to convert the description data into a form that is comparable to the classification scheme. Thus, we had to separate the description keywords within the full description and assign each trademark several keyword descriptions. This basically means transforming the data from Table 2 in the format that is shown in Table 4.

We therefore split the description to single lines and used common separators (semicolon, comma, slash 'and/or', 'and / or', 'and') as the respective cut-of lines. Next, we removed special characters, stopwords, other linking words of any relevance, and punctuations and set everything to lower case. We then lemmatized the descriptions with the 'textstem' package in 'R', so that the matching works with the lemmas of the respective words, i.e. the canonical, or the dictionary form of the word (e.g. 'ran' or 'running' is transformed to 'run'). Finally, we removed descriptions that are not related to the content of the trademark, like for example "or otherwise". Based on this cleaning, we run our matching algorithms.

Table 4: Lemmatized descriptions of trademarks

Application No.	Class	Goods/Services Description
17836735	12	Vehicle
17836735	12	Apparatus for locomot by land
17836735	12	Car
17836735	12	Part and fitting for vehicle
17833096	12	Motorized vehicle
17833096	12	rail vehicle
17833096	12	Motor
17833096	12	engine and drive for motor vehicle
17833096	12	Propulsion mechanism for automobil
17833096	12	Chassis for motor vehicle

Source: EUIPO

The aim of the matching procedure is to identify lemmas in the description of the applicant that match with the lemmas of the keywords from the explanatory note to the NICE classification or has a high similarity with it. For this purpose, we calculate the similarity between the respective lemmas at layer 5 of our classification within a given NICE class.

In the first step, the similarity value is calculated with the help of the Levenshtein distance, which calculates how many edits would be needed in order to align two text-strings. Edits can hereby be insertion, deletion or replacement of a character with another

one. The Levenshtein distance between the texts "motorized vehicle" and "Motorized vehicle", for example, would equal 1, replacement of the letter "s" by the letter "z". We then normalized the Levenshtein distance by the number of characters in a given string so it ranges between 0 and 1. We then subtract the distance measure from 1 to create a measure of similarity.

After obtaining the similarity scores, the entries to be selected as matches had to be determined. For this purpose, a predetermined threshold value  $t$  was calculated based on a manually created gold-standard of 200 entries (true matches). To determine the optimal threshold value  $t$  for the matching, we resorted to recall (ratio of correctly as positive identified elements in the total number of positive elements) and precision (ratio of correctly as positive identified elements in the total number of elements identified as positive) measures. The optimal compromise of the two concepts, i.e. the F-Score (harmonic mean of recall and precision) determines the  $t$ -value. The  $t$ -value was set to 0.89 on this basis.

If the similarity value of 0.89 is exceeded, the respective pair of lemmas is interpreted as a match. Since a trademark usually has several descriptive terms, multiple, parallel assignments of one trademark to several layer 5 classes are possible. The matching, however, is only performed within each NICE class, i.e. descriptive terms are only assigned to the classification within the same NICE class. If the similarity value exceeds the threshold value, the respective pair of entries is interpreted as a "match" and the trademark assigned to the classification keyword in question.

After this first matching step, we repeated the matching with a variant of the Levenshtein distance, namely the Jaro-Winkler distance. The Jaro-Winkler distance applies a scale to prefixes, thus giving a higher rating to strings that match from the beginning of the string. For this second step, a slightly higher  $t$ -value of 0.93 was assigned for the selection of the matches.

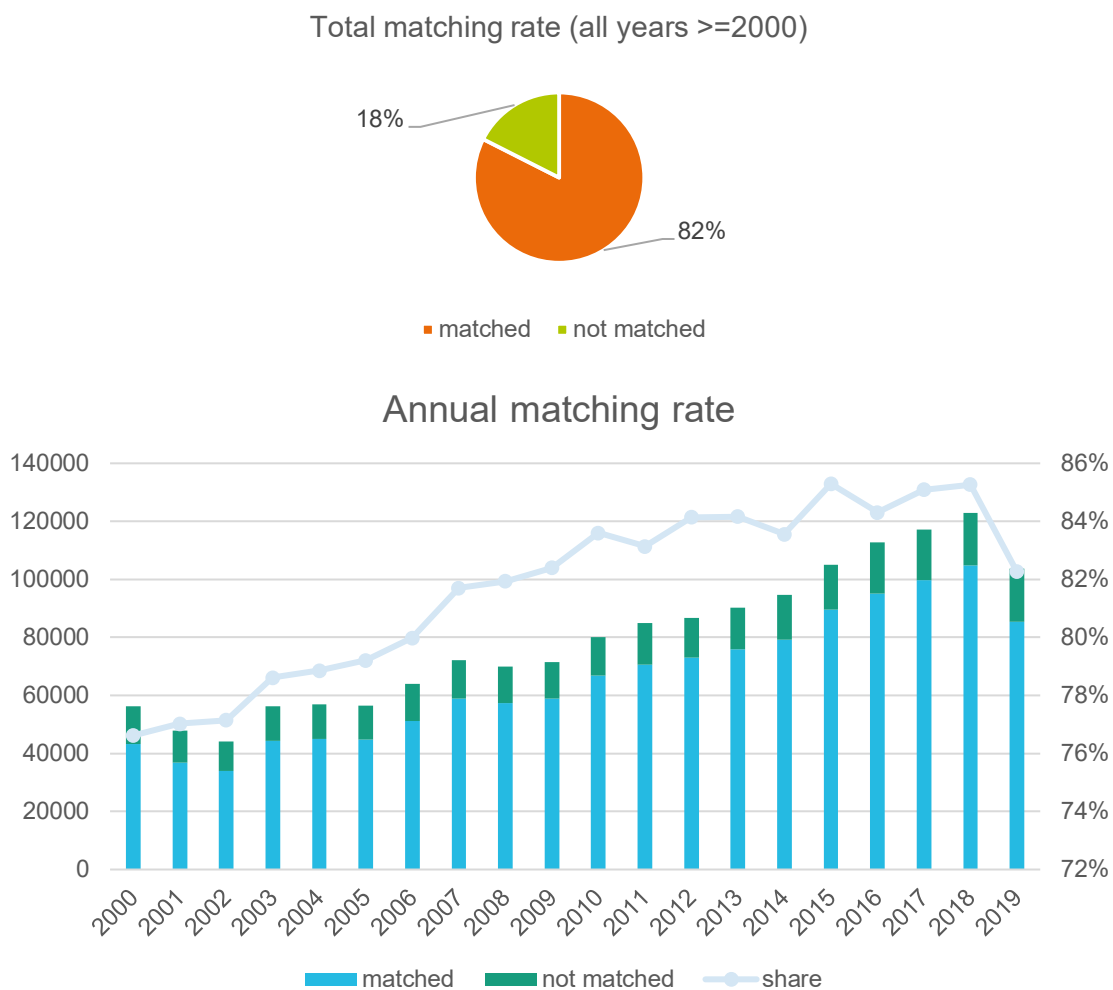
## **4 Results**

### **4.1 Coverage and validation of the matching**

The coverage of trademark filings in relation to all trademark filings at the EUIPO is plotted in Figure 2. In total, we reach a matching rate of 82%, i.e. we have assigned at least one classification code at level V to ~1.3 million trademarks out of ~1.6 million records in the EUIPO dataset from 2000 onwards. Applying the first step of the matching using the Levenshtein distance leads to a matching rate of 75% at a similarity ( $t$ -value) of 0.89. The second matching step using the Jaro-Winkler distance leads to an additional 7% of classification assignments.

When looking at the annual matching rate only minor differences can be found. The highest matching rate can be observed for 2018 (85%), which is the most recent year with complete data.<sup>3</sup> The matching rate drops rather constantly the further we go back in time, e.g. we only reach a matching rate of 77% in the year 2000. The explanation for this is that we use the explanatory note to the NICE classification of 2018. This list is constantly adapted to changes in (technical) language, implying that the 2018 version might include specific words that were not in use or phrases for technologies that as such did not exist in 2000.

Figure 2: Share of matched trademarks



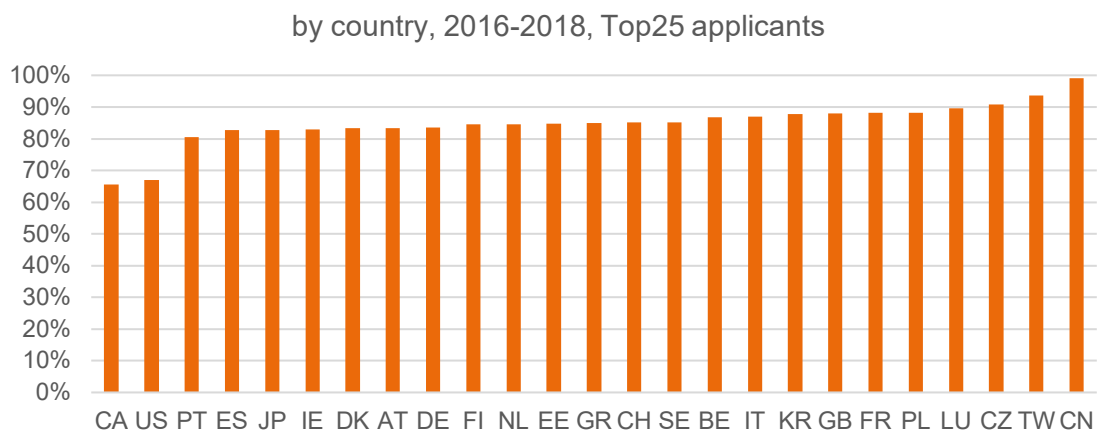
Source: EUIPO, own calculations

<sup>3</sup> We are using a snapshot of the EUIPO database version of autumn 2019 for the matching where the 2019 data is still incomplete-

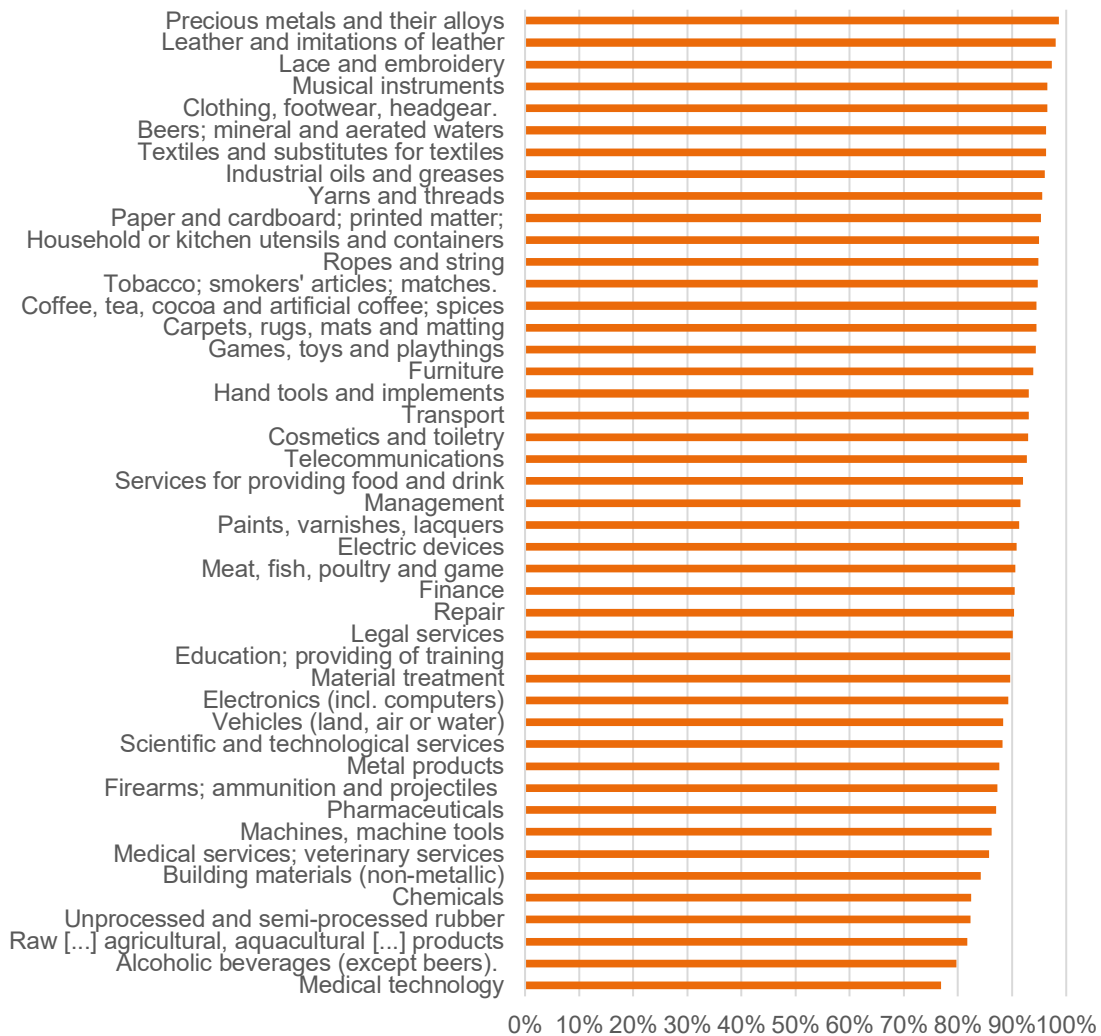


In order to validate our results, we can take a closer look at the share of matched trademarks by country and by NICE class to find out whether some countries/fields are over- or underrepresented in our matching. This is shown in Figure 3 for the period 2016 to 2018. As can be seen from the figure, the matching rate is similar across countries with the extremes being China with a matching rate of 99% on one end of the scale and Canada and the U.S. with 66% and 67%, respectively. All other countries are in the range of 81% and 94%. The explanation for this phenomenon is rooted in language differences. Applicants at the EUIPO can only provide their trademark descriptions in three languages, namely English, French and German. Consequently, the probability to choose from the pre-defined list of standardized keywords for the classification is much larger for non-native English, French or German speakers and especially the Chinese seem to rely heavily on the standardized keywords. For American and Canadian applicants, however, it is much easier to provide a description with self-provided keywords. This is aggravated further by the fact that the U.S. Patent and Trademark Office (USPTO) demands a detailed trademark description from applicants, most of which will not use the list of EUIPO keywords as reference. Since many trademark filings are simply translated when being filed at the EUIPO the standardized keywords are much less used by Americans and Canadians. Consequently, the differences between U.S./Canadian and Chinese applicants do not stem from systematic differences in our algorithm but in structural differences related to language and IPR systems. This has to be kept in mind for further analyses.

Figure 3: Share of matched trademarks by country and NICE class



## by NICE classes, 2016-2018



Source: EUIPO

When looking at the distribution of the matching rate by NICE class, it is found that the coverages range from 77% in medical technology (NICE class 10) to 99% in precious metals and alloys (NICE class 14). Thus, a rather high matching rate can be observed for all classes. When taking a look at the classes, it can be found that many classes related to chemistry and medicine have a somewhat lower coverage, while the largest matching rates can be found in the manufacturing of consumer goods. Apparently, the algorithm is working slightly better for discrete and less complex products and performs a bit worse for more complex products, even though no major distortions can be observed.

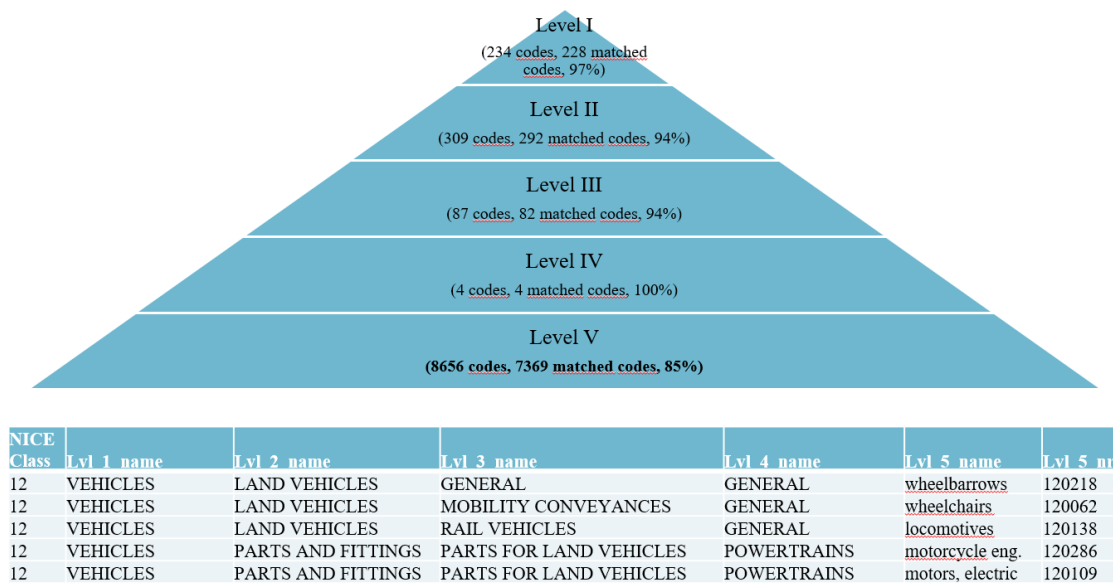
## 4.2 The classification setup

The classification is set up hierarchically with the NICE class being the top-level of the hierarchy. Below the NICE class, the classification has five levels, gradually getting more and more fine-grained. The hierarchy is taken over from the Explanatory Note to the NICE classification, which already provides a basic hierarchical setup. In case a classification does not exist for a certain level, we replaced it with a class called "General", to make sure that the same hierarchy level is available for each trademark. A trademark can be classified multiple times, i.e. there might be more than one level V class for any given trademark, also depending on how many NICE classes have been assigned to it upon filing.

An overview of the classification with an example for NICE class 12 ("vehicles and conveyances") is plotted in Figure 4. The least aggregated level, i.e. Level I, has 234 codes or entries. By means of the abovementioned methodology, 228 of these are assigned at least one trademark (97%). Some NICE classes are assigned more Level I classes than others. NICE class 11 ("Apparatus for lighting, heating, steam generating, cooking, refrigerating, ..."), for example, has 15 subclasses at Level I while class 23 ("Yarns and threads, for textile use" has only one Level I subclass. Having one level I subclass, however, only happens in three occasions (NICE classes 12, 23 and 38), while the average number of level I subclasses lies at 6 per NICE class.

From layer to layer, the classification becomes more and more fine-grained until the lowest level of aggregation, the Level V classification, is reached. Here, each class has a class number (level 5 number with 6 digits). The first two digits of this number represent the NICE class, while the remaining digits are more or less arbitrary. At Level V, there are potentially 8,656 classes of which 7,369 are assigned at least one trademark (85%).

Figure 4: The classification setup

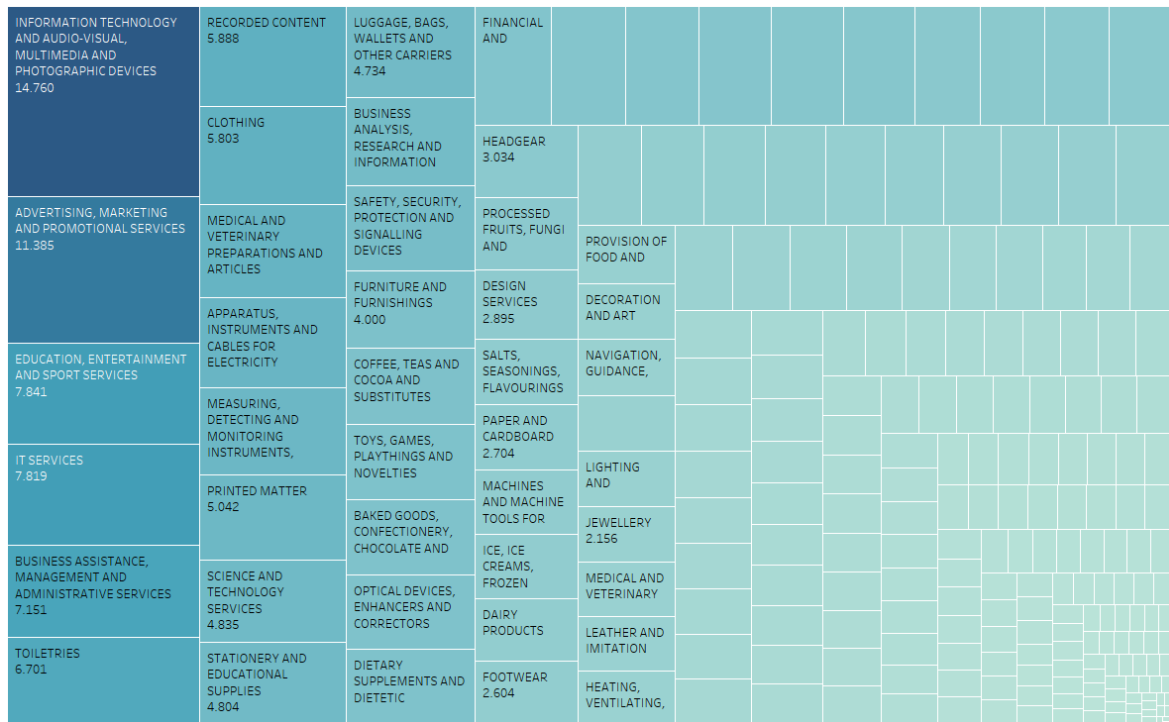


Source: EUIPO, WIPO, 2018

### 4.3 Examples and use cases

The level I classes by absolute filing figures for the complete EUIPO trademark landscape in 2018 is plotted in Figure 5. The largest level I class is "information technology and audio-visual, multimedia and photographic devices", a subclass of NICE class 9, "Electronics", followed by "advertising, marketing and promotional services", a subclass of NICE class 35 "Advertising; business management". The next largest level I classes worldwide are "education, entertainment and sports services", "IT services", "Business assistance, management and administrative services" and "toiletries", proving that the span across different brands is rather wide and the underlying technologies differ widely.

Figure 5: Trademark classification at Level I, World, 2018, absolute numbers

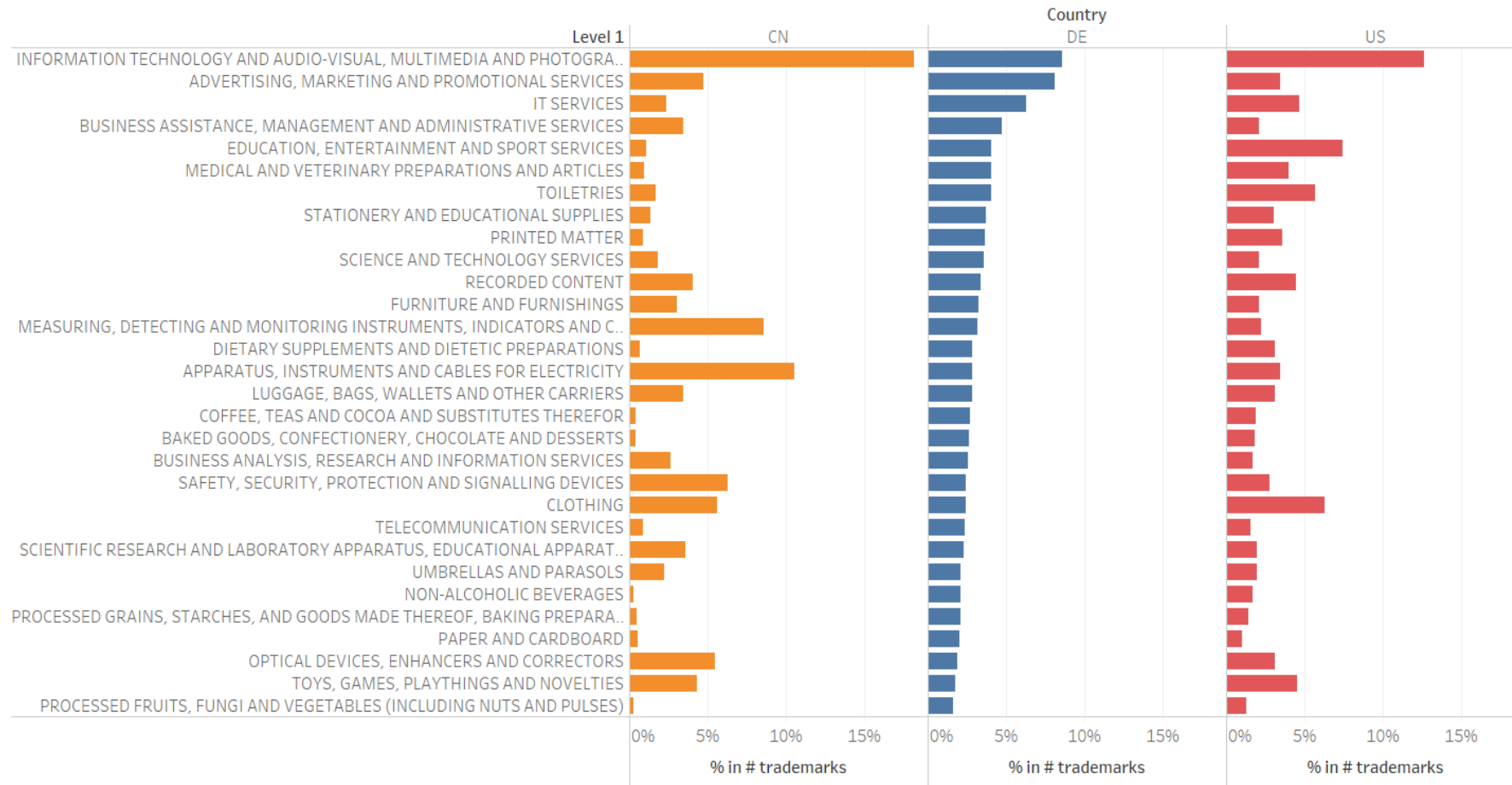


Source: EUIPO; calculations by Fraunhofer ISI

Note: This graph is used to give a graphical overview on the complete universe of Level I classes. The full classification can be found in the annex

The level I classification can also be differentiated by country (Figure 6). In the figure, the top thirty classes for Germany (as shares of total German filings) in comparison to the U.S. and China are plotted. As in the worldwide average, the largest level I class for all of the three countries is "information technology and audio-visual, multimedia and photographic devices". The next largest level I classes in Germany are "Advertising; marketing and promotional services", "IT services" and "Business assistance, management and administrative services". These classes have a much smaller weight in China, where especially "Apparatus, Instruments and Cables for Electricity", "Measuring, detecting and monitoring services", "safety, security, protection and signaling devices", "clothing" and "optical devices" have a much larger weight, once again showing that the Chinese trademark profile is more product than service oriented. For the U.S. the second largest level I class is "education and sports services", followed by "clothing", "toiletries" and "toys and games", i.e. classes that are more oriented to the consumer market than the "business market".

Figure 6: Shares of trademarks in total trademarks at classification level I, China, Germany and the U.S., 2018

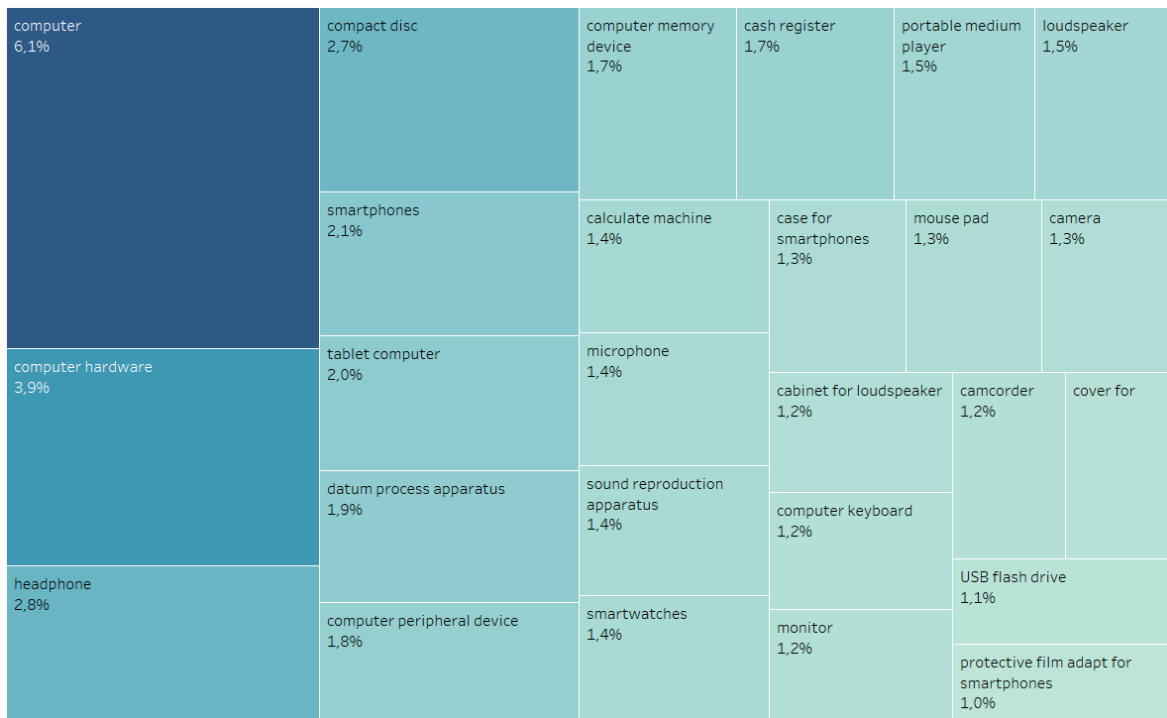


Source: EUIPO; calculations by Fraunhofer ISI

For the largest level I class, namely "information technology and audio-visual, multimedia and photographic devices" we have further generated an overview of the level V classes for the world (Figure 7) as well as Germany (Figure 8) in comparison. For the world, it can be found that the largest level V class in level I subclass "information technology and audio-visual, multimedia and photographic devices" is "computers", followed by "computer hardware", "headphones", "compact discs", "smartphones" and "tablet computers".

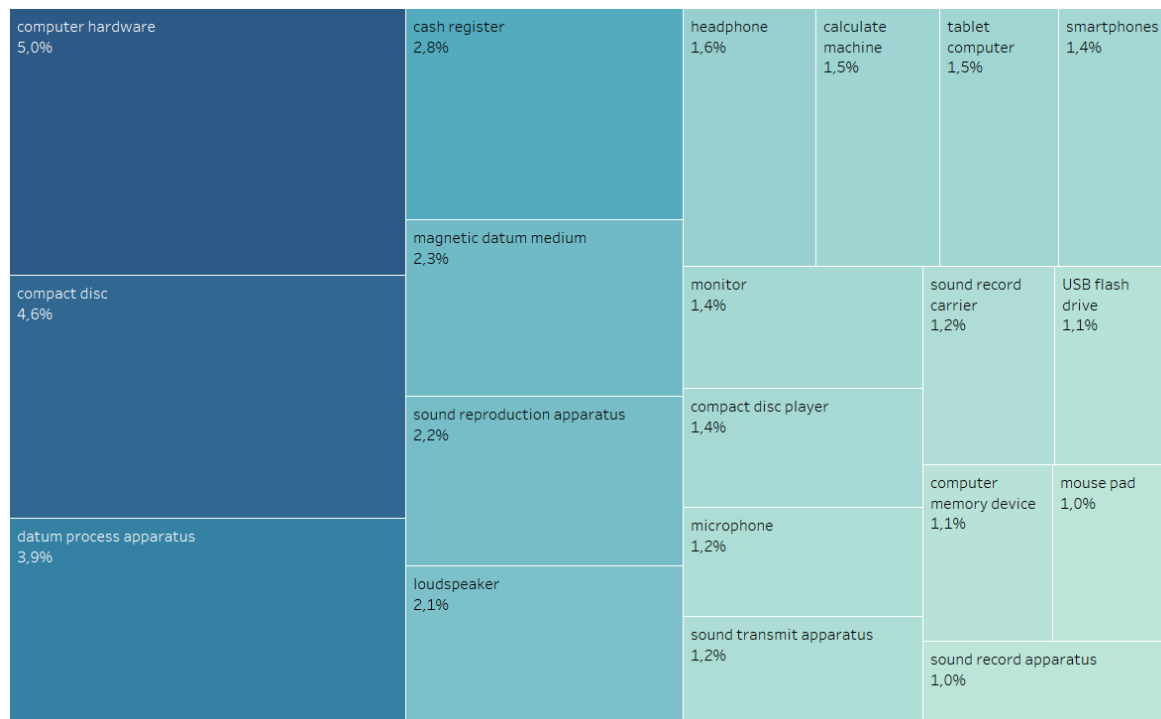
For Germany (Figure 8), in comparison, the largest level V classes in "information technology and audio-visual, multimedia and photographic devices" are "computer hardware", "compact disc", "datum process apparatus", "cash register", magnetic datum medium", "sound production apparatus" and "loudspeaker". This is a very good example of what differences can be found below the surface of the NICE classes and even below the level I subclasses as Germany shows a rather different picture than the world when it comes to this IT related level I subclass. It also shows the level of detail that can be reached with the help of the classification.

Figure 7: Shares of trademarks within level I class "information technology and audio-visual, multimedia and photographic devices" at level V, world, 2018



Source: EUIPO; calculations by Fraunhofer ISI

Figure 8: Shares of trademarks within level I class "information technology and audio-visual, multimedia and photographic devices" at level V, Germany, 2018



Source: EUIPO; calculations by Fraunhofer ISI

#### 4.4 Pitfalls

Although the classification works pretty well in generating trademark statistics at a detailed level including several hierarchies, there are still some pitfalls that are worth mentioning.

First and foremost, we have to acknowledge the fact that not all trademarks have been assigned to a classification entry. Still, 18% remain unclassified. As we have seen in the validation section, systematic distortions across countries and fields are limited. However, we still only deal with a subsample of all available trademark data. In addition, there are marks with several classification entries while there are others with only one classification. This is dependent on the length of the description provided by the applicant as well as - at least to a certain extent - the matching algorithm.

Second, hierarchy levels II, III and IV are not available for every trademark. At level IV, only one real classification code exists (namely "powertrains"). We have dealt with this issue by providing the sub-classification "General" for each trademark where a class is unavailable, to provide the most available detail possible - an option would have been to



leave out the sub-levels II to IV entirely. For analytical purposes, therefore, the intermediate sub-classes at levels II, III and IV should be used with caution - if at all. Different from the levels I and V, they do not (yet) form reliable stand-alone classifications, but rather auxiliary devices for those trying to understand the structure of the overall classification and to conceptual divide it into meaningful sub-areas.

Third, the hierarchy provided in the Explanatory Note to the NICE classification sometimes provides unclear results in the sense of a taxonomy. In the level I class "information technology and audio-visual, multimedia and photographic devices", for example (Figure 7), we can see that a level V class "computer hardware" exists alongside classes like "computer", "monitor" or "computer memory device". Here, it is not 100% clear whether these terms should exist in parallel or if it is the more generic term for the others.

Fourth, there are level V classes to which only one trademark is assigned, implying that this classes might not be "important" in the sense that this one assigned trademark might as well be assigned to another class, which would reduce the level of complexity. However, we kept all the available classes as there might be trademarks coming in in the future, which might be assigned to one of these classes with low filing figures at the moment. In addition, there are keywords or terms that occur very often, e.g. "musical instruments" in NICE class 15 ("musical instruments").

In sum, we can say that this classification provides a basic, automated structuring device for what is available in the data. We believe that this has its merits. In a first step, we decided to provide the raw data and leave the re-grouping of classes or re-assignments of specific trademarks to classes to manual, case by case decisions, in case an additional level of accuracy is needed for specific research efforts. We would argue, however, that our new method can even in those cases help to pre-structure the approach, and expedite the process. Despite all remaining shortcomings of the automated approach one needs to be aware that manual re-classification - while more precise - would at the same time mean that implicit interpretations enter into the interpretation of ambiguous cases.

## 5 Conclusions

In this paper, we have shown that classification of EUIPO trademarks can be set up on the basis of the available textual description that applicants provide for their trademarks during the filing process. The classification and the trademarks associated to it allow for quick and easy in-depth analyses, country and fields comparisons, time trends etc. without resorting to keyword-based analyses. It further allows us to look into much more thematic detail regarding EUIPO trademarks. This can also be seen as a stepping stone

for the prevalent question of how to differentiate knowledge-intensive from less knowledge-intensive trademarks which is an important distinction when it comes to the question of using trademarks as an indicator for innovative activities, especially in services, where broadly available indicators are scarce.

There still are open questions that could be subject to future research. We are especially interested in reaching a complete coverage of trademark data, i.e. we would like to assign all trademarks to at least one classification code. Here, a topic model or for NLP algorithms could help for the assignment of the remaining 18% of still unassigned trademarks. However, we are still at the experimental stage in this regard.

Another issue is the transferability of the classification to other patent offices in the world. Especially applying the method to the USPTO seems feasible. China could also be interesting, though we would then also have to deal with automated translations.

Finally, it remains to be seen whether it will prove feasible to keep five levels of classification below the NICE classes or whether it might eventually prove more informative to focus on only two levels. In any case, classifications are living constructs that constantly have to adapt to new developments, so we leave it to the pilot users of this classification to judge whether we (or others) will have to refine it in the future.

## 6 Appendix

In the following table, the classification at Level I is provided. The Level V classification is too large to be printed in the paper, but is available on the Fraunhofer ISI homepage at [https://www.isi.fraunhofer.de/content/dam/isi/dokumente/cci/innovation-systems-policy-analysis/2021/general\\_nice\\_classes\\_19o.xlsx](https://www.isi.fraunhofer.de/content/dam/isi/dokumente/cci/innovation-systems-policy-analysis/2021/general_nice_classes_19o.xlsx). Fraunhofer ISI hosts the EUIPO (as well as USPTO) trademark data as a relational database (Oracle relational database system (RDBMS)) within the RISIS2 (Research Infrastructure for Science and Innovation Policy studies, <https://www.risis2.eu/>) project. Researchers can apply for using the data via the RISIS data infrastructure and will then also get direct access to the classification, i.e. the assignment of trademark application numbers to the classes at all levels.

Table 5: The trademark classification at level I

NICE code	NICE name	LEVEL I name
1	Chemicals	UNPROCESSED PLASTICS
1	Chemicals	ADHESIVES FOR USE IN INDUSTRY
1	Chemicals	GROWING MEDIA, FERTILIZERS AND CHEMICALS FOR USE IN AGRICULTURE, HORTICULTURE AND FORESTRY
1	Chemicals	DETERGENTS FOR USE IN MANUFACTURE AND INDUSTRY
1	Chemicals	CHEMICAL SUBSTANCES, CHEMICAL MATERIALS AND CHEMICAL PREPARATIONS, AND NATURAL ELEMENTS
1	Chemicals	FILTERING MATERIALS [CHEMICAL, MINERAL, VEGETABLE AND OTHER UNPROCESSED MATERIALS]
1	Chemicals	CHEMICAL PREPARATIONS AND MATERIALS FOR FILM, PHOTOGRAPHY AND PRINTING
1	Chemicals	CHEMICAL AND ORGANIC COMPOSITIONS FOR USE IN THE MANUFACTURE OF FOOD AND BEVERAGES
1	Chemicals	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
1	Chemicals	SALTS FOR INDUSTRIAL PURPOSES
1	Chemicals	UNPROCESSED ARTIFICIAL AND SYNTHETIC RESINS
1	Chemicals	STARCHES FOR USE IN MANUFACTURING AND INDUSTRY
1	Chemicals	PUTTIES, AND FILLERS AND PASTES FOR USE IN INDUSTRY
2	Paints, varnishes, lacquers	DYES, COLORANTS, PIGMENTS AND INKS
2	Paints, varnishes, lacquers	THINNERS AND THICKENERS FOR COATINGS, DYES AND INKS
2	Paints, varnishes, lacquers	COATINGS
2	Paints, varnishes, lacquers	RAW NATURAL RESINS
2	Paints, varnishes, lacquers	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
3	Cosmetics and toiletry	CLEANING AND FRAGRANCING PREPARATIONS
3	Cosmetics and toiletry	TOILETRIES
3	Cosmetics and toiletry	ANIMAL GROOMING PREPARATIONS
3	Cosmetics and toiletry	ESSENTIAL OILS AND AROMATIC EXTRACTS

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
3	Cosmetics and toiletry	ABRASIVES
3	Cosmetics and toiletry	TAILORS AND COBBLERS WAX
3	Cosmetics and toiletry	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
4	Industrial oils and greases	FUELS AND ILLUMINANTS
4	Industrial oils and greases	DUST CONTROLLING COMPOSITIONS
4	Industrial oils and greases	LUBRICANTS AND INDUSTRIAL GREASES, WAXES AND FLUIDS
4	Industrial oils and greases	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
5	Pharmaceuticals	DIETARY SUPPLEMENTS AND DIETETIC PREPARATIONS
5	Pharmaceuticals	DENTAL PREPARATIONS AND ARTICLES, AND MEDICATED DENTIFRICES
5	Pharmaceuticals	SANITARY PREPARATIONS AND ARTICLES
5	Pharmaceuticals	MEDICAL AND VETERINARY PREPARATIONS AND ARTICLES
5	Pharmaceuticals	PEST CONTROL PREPARATIONS AND ARTICLES
6	Metal products	UNPROCESSED AND SEMI-PROCESSED MATERIALS OF METAL, NOT SPECIFIED FOR USE
6	Metal products	METAL HARDWARE
6	Metal products	BUILDING AND CONSTRUCTION MATERIALS AND ELEMENTS OF METAL
6	Metal products	DOORS, GATES, WINDOWS AND WINDOW COVERINGS OF METAL
6	Metal products	STRUCTURES AND TRANSPORTABLE BUILDINGS OF METAL
6	Metal products	STATUES AND WORKS OF ART OF COMMON METALS
6	Metal products	CONTAINERS, AND TRANSPORTATION AND PACKAGING ARTICLES, OF METAL
6	Metal products	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
7	Machines, machine tools	MOVING AND HANDLING EQUIPMENT
7	Machines, machine tools	MACHINES AND MACHINE TOOLS FOR TREATMENT OF MATERIALS AND FOR MANUFACTURING
7	Machines, machine tools	AGRICULTURAL, EARTHMOVING, CONSTRUCTION, OIL AND GAS EXTRACTION AND MINING EQUIPMENT

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
7	Machines, machine tools	PUMPS, COMPRESSORS AND FANS
7	Machines, machine tools	GENERATORS OF ELECTRICITY
7	Machines, machine tools	ENGINES, POWERTRAINS, AND MACHINE PARTS, AND CONTROLS FOR THE OPERATION OF MACHINES AND ENGINES
7	Machines, machine tools	DISPENSING MACHINES
7	Machines, machine tools	SWEEPING, CLEANING, WASHING AND LAUNDERING MACHINES
7	Machines, machine tools	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
8	Hand tools and implements	HAND-OPERATED TOOLS AND IMPLEMENTS FOR TREATMENT OF MATERIALS, AND FOR CONSTRUCTION, REPAIR AND MAINTENANCE
8	Hand tools and implements	HYGIENIC AND BEAUTY IMPLEMENTS FOR HUMANS AND ANIMALS
8	Hand tools and implements	EDGED AND BLUNT WEAPONS
8	Hand tools and implements	FOOD PREPARATION IMPLEMENTS, KITCHEN KNIVES AND CUTTING IMPLEMENTS, CUTLERY FOR EATING
8	Hand tools and implements	LIFTING TOOLS
9	Electronics (incl. computers)	RECORDED CONTENT
9	Electronics (incl. computers)	INFORMATION TECHNOLOGY AND AUDIO-VISUAL, MULTIMEDIA AND PHOTOGRAPHIC DEVICES
9	Electronics (incl. computers)	OPTICAL DEVICES, ENHANCERS AND CORRECTORS
9	Electronics (incl. computers)	SAFETY, SECURITY, PROTECTION AND SIGNALLING DEVICES
9	Electronics (incl. computers)	MAGNETS, MAGNETIZERS AND DEMAGNETIZERS
9	Electronics (incl. computers)	SCIENTIFIC AND LABORATORY DEVICES FOR TREATMENT USING ELECTRICITY

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
9	Electronics (incl. computers)	APPARATUS, INSTRUMENTS AND CABLES FOR ELECTRICITY
9	Electronics (incl. computers)	MEASURING, DETECTING AND MONITORING INSTRUMENTS, INDICATORS AND CONTROLLERS
9	Electronics (incl. computers)	SCIENTIFIC RESEARCH AND LABORATORY APPARATUS, EDUCATIONAL APPARATUS AND SIMULATORS
9	Electronics (incl. computers)	NAVIGATION, GUIDANCE, TRACKING, TARGETING AND MAP MAKING DEVICES
9	Electronics (incl. computers)	DIVING EQUIPMENT
10	Medical technology	PHYSICAL THERAPY EQUIPMENT
10	Medical technology	HEARING PROTECTION DEVICES
10	Medical technology	FEEDING AIDS AND PACIFIERS
10	Medical technology	SEX AIDS
10	Medical technology	MEDICAL AND VETERINARY APPARATUS AND INSTRUMENTS
10	Medical technology	ORTHOPEDIC AND MOBILITY AIDS
10	Medical technology	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
10	Medical technology	MEDICAL FURNITURE AND BEDDING, EQUIPMENT FOR MOVING PATIENTS
10	Medical technology	MEDICAL CLOTHING
10	Medical technology	PROSTHETICS AND ARTIFICIAL IMPLANTS
11	Electric devices	SANITARY INSTALLATIONS, WATER SUPPLY AND SANITATION EQUIPMENT
11	Electric devices	BURNERS, BOILERS AND HEATERS
11	Electric devices	LIGHTING AND LIGHTING REFLECTORS
11	Electric devices	FOOD AND BEVERAGE COOKING, HEATING, COOLING AND TREATMENT EQUIPMENT
11	Electric devices	FLUES AND INSTALLATIONS FOR CONVEYING EXHAUST GASES
11	Electric devices	SUN TANNING APPLIANCES
11	Electric devices	HEATING, VENTILATING, AND AIR CONDITIONING AND PURIFICATION EQUIPMENT (AMBIENT)
11	Electric devices	IGNITERS
11	Electric devices	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
11	Electric devices	FIREPLACES
11	Electric devices	FILTERS FOR INDUSTRIAL AND HOUSEHOLD USE
11	Electric devices	INDUSTRIAL TREATMENT INSTALLATIONS

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
11	Electric devices	PERSONAL HEATING AND DRYING IMPLEMENTS
11	Electric devices	DRYING INSTALLATIONS
11	Electric devices	REFRIGERATING AND FREEZING EQUIPMENT
11	Electric devices	REGULATING AND SAFETY ACCESSORIES FOR WATER AND GAS INSTALLATIONS
11	Electric devices	NUCLEAR INSTALLATIONS
12	Vehicles (land, air or water)	VEHICLES AND CONVEYANCES
13	Firearms; ammunition and projectiles	EXPLOSIVE SUBSTANCES AND DEVICES, OTHER THAN ARMS
13	Firearms; ammunition and projectiles	WEAPONS AND AMMUNITION
14	Precious metals and their alloys	JEWELLERY
14	Precious metals and their alloys	TIME INSTRUMENTS
14	Precious metals and their alloys	OTHER ARTICLES OF PRECIOUS METALS AND PRECIOUS STONES, AND IMITATIONS THEREOF
14	Precious metals and their alloys	JEWELLERY BOXES AND WATCH BOXES
14	Precious metals and their alloys	KEY RINGS AND KEY CHAINS, AND CHARMS THEREFOR
14	Precious metals and their alloys	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
14	Precious metals and their alloys	GEMSTONES, PEARLS AND PRECIOUS METALS, AND IMITATIONS THEREOF
15	Musical instruments	MUSICAL INSTRUMENTS
15	Musical instruments	MUSICAL ACCESSORIES
15	Musical instruments	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
16	Paper and cardboard; printed matter;	BAGS AND ARTICLES FOR PACKAGING, WRAPPING AND STORAGE OF PAPER, CARDBOARD OR PLASTICS
16	Paper and cardboard; printed matter;	STATIONERY AND EDUCATIONAL SUPPLIES

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
16	Paper and cardboard; printed matter;	WORKS OF ART AND FIGURINES OF PAPER AND CARDBOARD, AND ARCHITECTS MODELS
16	Paper and cardboard; printed matter;	DECORATION AND ART MATERIALS AND MEDIA
16	Paper and cardboard; printed matter;	FILTERING MATERIALS OF PAPER
16	Paper and cardboard; printed matter;	PAPER AND CARDBOARD
16	Paper and cardboard; printed matter;	MONEY HOLDERS
16	Paper and cardboard; printed matter;	DISPOSABLE PAPER PRODUCTS
16	Paper and cardboard; printed matter;	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
16	Paper and cardboard; printed matter;	ADHESIVES FOR STATIONERY OR HOUSEHOLD PURPOSES
16	Paper and cardboard; printed matter;	PRINTED MATTER
17	Unprocessed and semi-processed rubber	FLEXIBLE PIPES, TUBES, HOSES AND FITTINGS THEREFOR (INCLUDING VALVES), AND FITTINGS FOR RIGID PIPES, ALL NON-METALLIC
17	Unprocessed and semi-processed rubber	UNPROCESSED AND SEMI-PROCESSED MATERIALS INCLUDED IN THE CLASS, NOT SPECIFIED FOR USE



<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
17	Unprocessed and semi-processed rubber	FINISHED OR SEMI-FINISHED GOODS, OF MATERIALS INCLUDED IN THE CLASS, SPECIFIED FOR USE
17	Unprocessed and semi-processed rubber	SEALS, SEALANTS AND FILLERS
17	Unprocessed and semi-processed rubber	INSULATION AND BARRIER ARTICLES AND MATERIALS
18	Leather and imitations of leather	UMBRELLAS AND PARASOLS
18	Leather and imitations of leather	WALKING STICKS
18	Leather and imitations of leather	LUGGAGE, BAGS, WALLETS AND OTHER CARRIERS
18	Leather and imitations of leather	LEATHER AND IMITATION LEATHER, PELTS AND HIDES, AND GOODS MADE THEREOF
18	Leather and imitations of leather	SADDLERY, WHIPS AND APPAREL FOR ANIMALS
18	Leather and imitations of leather	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
19	Building materials (non-metallic)	BUILDING AND CONSTRUCTION MATERIALS AND ELEMENTS, NOT OF METAL
19	Building materials (non-metallic)	DOORS, GATES, WINDOWS AND WINDOW COVERINGS, NOT OF METAL
19	Building materials (non-metallic)	STRUCTURES AND TRANSPORTABLE BUILDINGS, NOT OF METAL
19	Building materials (non-metallic)	UNPROCESSED AND SEMI-PROCESSED MATERIALS INCLUDED IN THE CLASS, NOT SPECIFIED FOR USE
19	Building materials (non-metallic)	STATUES AND WORKS OF ART MADE OF MATERIALS SUCH AS STONE, CONCRETE AND MARBLE, INCLUDED IN THE CLASS
20	Furniture	NON-METALLIC HARDWARE
20	Furniture	STATUES, FIGURINES, WORKS OF ART AND ORNAMENTS AND DECORATIONS, MADE OF MATERIALS SUCH AS WOOD, WAX, PLASTER OR PLASTIC, INCLUDED IN THE CLASS

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
20	Furniture	FURNITURE AND FURNISHINGS
20	Furniture	CONTAINERS, AND CLOSURES AND HOLDERS THEREFOR, NON-METALLIC
20	Furniture	DISPLAYS, STANDS AND SIGNAGE, NON-METALLIC
20	Furniture	ANIMAL HOUSING AND BEDS
20	Furniture	UNPROCESSED AND SEMI-PROCESSED MATERIALS INCLUDED IN THE CLASS, NOT SPECIFIED FOR USE
20	Furniture	LADDERS AND MOVABLE STEPS, NON-METALLIC
21	Household or kitchen utensils and containers	UNWORKED AND SEMI-WORKED GLASS, NOT SPECIFIED FOR USE
21	Household or kitchen utensils and containers	GARDENING ARTICLES
21	Household or kitchen utensils and containers	BRUSHES AND OTHER ARTICLES FOR CLEANING, BRUSH-MAKING MATERIALS
21	Household or kitchen utensils and containers	STATUES, FIGURINES, PLAQUES AND WORKS OF ART, MADE OF MATERIALS SUCH AS PORCELAIN, TERRA-COTTA OR GLASS, INCLUDED IN THE CLASS
21	Household or kitchen utensils and containers	TABLEWARE, COOKWARE AND CONTAINERS
21	Household or kitchen utensils and containers	COSMETIC AND TOILET UTENSILS AND BATHROOM ARTICLES
21	Household or kitchen utensils and containers	ARTICLES FOR ANIMALS
21	Household or kitchen utensils and containers	ARTICLES FOR THE CARE OF CLOTHING AND FOOTWEAR
22	Ropes and string	GOODS MADE OF TEXTILE AND FIBERS
22	Ropes and string	PADDING AND STUFFING MATERIALS
22	Ropes and string	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
22	Ropes and string	RAW TEXTILE FIBERS AND SUBSTITUTES
23	Yarns and threads	YARNS AND THREADS
24	Textiles and substitutes for textiles	FABRICS
24	Textiles and substitutes for textiles	TEXTILE GOODS, AND SUBSTITUTES FOR TEXTILE GOODS
24	Textiles and substitutes for textiles	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
25	Clothing, footwear, headgear.	HEADGEAR
25	Clothing, footwear, headgear.	CLOTHING
25	Clothing, footwear, headgear.	FOOTWEAR
25	Clothing, footwear, headgear.	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
26	Lace and embroidery	ACCESSORIES FOR APPAREL, SEWING ARTICLES AND DECORATIVE TEXTILE ARTICLES
26	Lace and embroidery	HAIR ORNAMENTS, HAIR ROLLERS, HAIR FASTENING ARTICLES, AND FALSE HAIR
26	Lace and embroidery	ARTIFICIAL FRUIT, FLOWERS AND VEGETABLES
26	Lace and embroidery	NEEDLES AND PINS FOR ENTOMOLOGY
26	Lace and embroidery	CHARMS [NOT JEWELLERY OR FOR KEYS, RINGS OR CHAINS]
27	Carpets, rugs, mats and matting	FLOOR COVERINGS AND ARTIFICIAL GROUND COVERINGS
27	Carpets, rugs, mats and matting	WALL AND CEILING COVERINGS
28	Games, toys and playthings	SPORTING ARTICLES AND EQUIPMENT
28	Games, toys and playthings	FESTIVE DECORATIONS AND ARTIFICIAL CHRISTMAS TREES

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
28	Games, toys and playthings	TOYS, GAMES, PLAYTHINGS AND NOVELTIES
28	Games, toys and playthings	FAIRGROUND AND PLAYGROUND APPARATUS
29	Meat, fish, poultry and game	PROCESSED FRUITS, FUNGI AND VEGETABLES (INCLUDING NUTS AND PULSES)
29	Meat, fish, poultry and game	SOUPS AND STOCKS, MEAT EXTRACTS
29	Meat, fish, poultry and game	PREPARED INSECTS AND LARVAE
29	Meat, fish, poultry and game	SAUSAGE SKINS AND IMITATIONS THEREOF
29	Meat, fish, poultry and game	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
29	Meat, fish, poultry and game	MEATS
29	Meat, fish, poultry and game	FISH, SEAFOOD AND MOLLUSCS
29	Meat, fish, poultry and game	DAIRY PRODUCTS AND DAIRY SUBSTITUTES
29	Meat, fish, poultry and game	BIRDS EGGS AND EGG PRODUCTS
29	Meat, fish, poultry and game	OILS AND FATS
30	Coffee, tea, cocoa and artificial coffee; spices	SUGARS, NATURAL SWEETENERS, SWEET COATINGS AND FILLINGS, BEE PRODUCTS
30	Coffee, tea, cocoa and artificial coffee; spices	BAKED GOODS, CONFECTIONERY, CHOCOLATE AND DESSERTS
30	Coffee, tea, cocoa and artificial coffee; spices	ICE, ICE CREAMS, FROZEN YOGURTS AND SORBETS

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
30	Coffee, tea, cocoa and artificial coffee; spices	COFFEE, TEAS AND COCOA AND SUBSTITUTES THEREFOR
30	Coffee, tea, cocoa and artificial coffee; spices	PROCESSED GRAINS, STARCHES, AND GOODS MADE THEREOF, BAKING PREPARATIONS AND YEASTS
30	Coffee, tea, cocoa and artificial coffee; spices	CONVENIENCE FOOD AND SAVORY SNACKS
30	Coffee, tea, cocoa and artificial coffee; spices	SALTS, SEASONINGS, FLAVOURINGS AND CONDIMENTS
31	Raw [...] agricultural, aquacultural [...] products	FOODSTUFFS AND FODDER FOR ANIMALS
31	Raw [...] agricultural, aquacultural [...] products	BEDDING AND LITTER FOR ANIMALS
31	Raw [...] agricultural, aquacultural [...] products	AGRICULTURAL AND AQUACULTURAL CROPS, HORTICULTURE AND FORESTRY PRODUCTS
31	Raw [...] agricultural, aquacultural [...] products	LIVE ANIMALS, ORGANISMS FOR BREEDING
32	Beers; mineral and aerated waters	BEER AND BREWERY PRODUCTS
32	Beers; mineral and aerated waters	NON-ALCOHOLIC BEVERAGES
32	Beers; mineral and aerated waters	PREPARATIONS FOR MAKING BEVERAGES
32	Beers; mineral and aerated waters	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
33	Alcoholic beverages (except beers).	PREPARATIONS FOR MAKING ALCOHOLIC BEVERAGES

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
33	Alcoholic beverages (except beers).	ALCOHOLIC BEVERAGES (EXCEPT BEER)
33	Alcoholic beverages (except beers).	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
34	Tobacco; smokers' articles; matches.	MATCHES
34	Tobacco; smokers' articles; matches.	TOBACCO AND TOBACCO PRODUCTS (INCLUDING SUBSTITUTES)
34	Tobacco; smokers' articles; matches.	ARTICLES FOR USE WITH TOBACCO
34	Tobacco; smokers' articles; matches.	PERSONAL VAPORISERS AND ELECTRONIC CIGARETTES, AND FLAVOURINGS AND SOLUTIONS THEREFOR
34	Tobacco; smokers' articles; matches.	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
35	Management	ADVERTISING, MARKETING AND PROMOTIONAL SERVICES
35	Management	BUSINESS ASSISTANCE, MANAGEMENT AND ADMINISTRATIVE SERVICES
35	Management	BUSINESS ANALYSIS, RESEARCH AND INFORMATION SERVICES
35	Management	COMMERCIAL TRADING AND CONSUMER INFORMATION SERVICES
36	Finance	INSURANCE SERVICES
36	Finance	REAL ESTATE SERVICES
36	Finance	PAWNBROKERAGE
36	Finance	PROVISION OF PREPAID CARDS AND TOKENS
36	Finance	SAFE DEPOSIT SERVICES
36	Finance	FINANCIAL AND MONETARY SERVICES, AND BANKING
36	Finance	FINANCIAL AND MONETARY SERVICES, AND BANKING
36	Finance	FUNDRAISING AND FINANCIAL SPONSORSHIP
36	Finance	VALUATION SERVICES
36	Finance	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
37	Repair	INSTALLATION, CLEANING, REPAIR AND MAINTENANCE
37	Repair	BUILDING, CONSTRUCTION AND DEMOLITION
37	Repair	RENTAL OF TOOLS, PLANT AND EQUIPMENT FOR CONSTRUCTION AND DEMOLITION
37	Repair	EXTRACTION OF NATURAL RESOURCES
37	Repair	EXTERMINATION, DISINFECTION AND PEST CONTROL

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
38	Telecommunications	TELECOMMUNICATION SERVICES
39	Transport	TRANSPORT
39	Transport	PACKAGING AND STORAGE OF GOODS
39	Transport	VEHICLE PARKING AND STORAGE
39	Transport	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
39	Transport	DISTRIBUTION BY PIPELINE AND CABLE
40	Material treatment	CUSTOM MANUFACTURE AND ASSEMBLY SERVICES
40	Material treatment	ENERGY PRODUCTION
40	Material treatment	TREATMENT AND TRANSFORMATION OF MATERIALS
40	Material treatment	PRINTING, AND PHOTOGRAPHIC AND CINEMATOGRAPHIC DEVELOPMENT
40	Material treatment	AIR AND WATER CONDITIONING AND PURIFICATION
40	Material treatment	SLAUGHTERING
40	Material treatment	FOOD AND BEVERAGE TREATMENT
40	Material treatment	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
41	Education; providing of training	EDUCATION, ENTERTAINMENT AND SPORT SERVICES
41	Education; providing of training	PUBLISHING AND REPORTING
41	Education; providing of training	TRANSLATION AND INTERPRETATION
41	Education; providing of training	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
42	Scientific and technological services	IT SERVICES
42	Scientific and technological services	SCIENCE AND TECHNOLOGY SERVICES
42	Scientific and technological services	TESTING, AUTHENTICATION AND QUALITY CONTROL
42	Scientific and technological services	DESIGN SERVICES
43	Services for providing food and drink	TEMPORARY ACCOMMODATION

<b>NICE code</b>	<b>NICE name</b>	<b>LEVEL I name</b>
43	Services for providing food and drink	RENTAL OF FURNITURE, LINENS AND TABLE SETTINGS
43	Services for providing food and drink	PROVISION OF FOOD AND DRINK
43	Services for providing food and drink	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
44	Medical services; veterinary services	AGRICULTURE, AQUACULTURE, HORTICULTURE AND FORESTRY SERVICES
44	Medical services; veterinary services	HUMAN HEALTHCARE SERVICES
44	Medical services; veterinary services	HUMAN HYGIENE AND BEAUTY CARE
44	Medical services; veterinary services	GOODS OR SERVICES NOT ASSOCIATED WITH A HIERARCHY GROUP
44	Medical services; veterinary services	ANIMAL HEALTHCARE SERVICES
45	Legal services	PERSONAL AND SOCIAL SERVICES
45	Legal services	LEGAL SERVICES
45	Legal services	SAFETY, RESCUE, SECURITY AND ENFORCEMENT SERVICES



## 7 References

- Castaldi, C., 2019. All the great things you can do with trademark data: Taking stock and looking ahead. *Strategic Organization*.
- deGrazia, C., Myers, A., Toole, A.A., 2020. Innovation activities and business cycles: are trademarks a leading indicator? *Industry and Innovation* 27 (1-2), 184–203.
- Flikkema, M., Castaldi, C., Man, A.-P. de, Seip, M., 2019. Trademarks' relatedness to product and service innovation: A branding strategy approach. *Research Policy* 48 (6), 1340–1353.
- Flikkema, M., Man, A.-P. de, Castaldi, C., 2014. Are Trademark Counts a Valid Indicator of Innovation? Results of an In-Depth Study of New Benelux Trademarks Filed by SMEs. *Industry and Innovation* 21 (4), 310–331.
- Gauch, S., 2007. Marken als Innovationsindikator. Studien zum deutschen Innovations-system Nr. 12-2007. Fraunhofer-Institut für System- und Innovationsforschung ISI, Karlsruhe.
- Greenhalgh, C., Rogers, M., 2006. The value of innovation: The interaction of competi-tion, R&D and IP. *Research Policy* 35 (4), 562–580.
- Krishnapuram, R., Kummamuru, K., 2003. Automatic Taxonomy Generation: Issues and Possibilities, in: *Fuzzy Sets and Systems — IFSA 2003*, Berlin, Heidelberg. 2003. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 52–63.
- Mendonca, S., Pereira, T.S., Godinho, M.M., 2004. Trademarks as an Indicator of Inno-vation and Industrial Change. LEM Working Paper Series No. 2004/15, Pisa.
- Navarro, G., 2001. A guided tour to approximate string matching. *ACM Computing Sur-veys* 33 (1), 31–88.
- Perepechko, L.N., 2018. Trademarks protection and development of industrial produc-tion, in: . Proceedings of the 32nd International Business Information Management Association Conference, IBIMA 2018 - Vision 2020: Sustainable Economic Devel-opment and Application of Innovation Management from Regional expansion to Global Growth, pp. 231–239.
- Sandner, P.G., Block, J., 2011. The market value of R&D, patents, and trademarks. *Re-search Policy* 40, 969–985.
- Schmoch, U., 2003. Service marks as novel innovation indicator. *Research Evaluation* 12 (2), 149–156.

- Schmoch, U., 2014. Knowledge transfer from German universities into the service sector as reflected by service marks. *Research Evaluation* 23 (2014), 341–351.
- Schmoch, U., Gauch, S., 2009. Service marks as indicators for innovation in knowledge-based services. *Research Evaluation* 18 (4), 323–335.
- Thoma, G., 2019. Composite value index of trademark indicators. *World Patent Information* 56, 64–75.
- Winkler, W.E., 1990. String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage. *Proceedings of the Section on Survey Research Methods - American Statistical Association*, 354–359.
- WIPO, 2018. List of Goods and Services with Class Headings and Explanatory Notes in PDF format.
- Yang, L., Yin, X., He, Y., 2018. Trademark protection, quality improvement and exports in developing countries. *World Economy* 41 (11), 3242–3254.