Blockchain patent landscaping: an expert based methodology and search query

Nigel S. Clarke [a,*], Björn Jürgens [b], Victor Herrero-Solana [c]

[a] *European Patent Office, Patent Information Research, Vienna, Austria*
[b] *Agency of Innovation and Development of Andalusia, CITPIA Patlib Centre, Seville, Spain*
[c] *University of Granada, SCImago-UGR (SEJ036), Granada, Spain*

**Corresponding author. E-mail addresses:* nclarke@epo.org (N.S. Clarke), bjurgens@agenciaidea.es (B. Jürgens), victorhs@ugr.es (V. Herrero-Solana).

World Patent Information 61 (2020) 101964

https://doi.org/10.1016/j.wpi.2020.101964

Abstract

The present study is concerned with the emergence of Blockchain related technologies in terms of patenting activity. Blockchain has captured the attention of the public and research has intensified in this field over the last few years, making it a highly interesting topic of study for a patent analysis in order to obtain insight into the developments of this emerging technology. In this paper we present a unique methodology and exhaustive search strategy for identifying Blockchain patent documents by using a combination of specific keywords and patent classifications. This query was built in cooperation with subject matter experts of the European Patent Office (EPO). Our keyword set was then analysed by relevance and was prioritised. The set of specific relevant patent classifications was furthermore combined with keywords in order to exclude false positives. With our methodology we present an exhaustive query for retrieving a highly relevant dataset of Blockchain related patents, extracted from the EPO databases that can be used for patent landscaping exercises or any other bibliometric analysis. In a case study we applied the search strategy to analyse worldwide Blockchain patenting from 2008 till 2018.

Keywords:

blockchain; search; query; patent analysis; patent landscape; european patent office

1. Introduction

Blockchains are transparent and decentralised methods of recording lists of transactions and have become famous as the technology behind digital currencies such as Bitcoin. In general, there is often confusion with these terms and they are often used interchangeably to indicate some of the three main concepts they are based on: the underlying Blockchain technology, the protocol and the client, and the cryptocurrency itself [1].

The idea of Bitcoin and Blockchain was born in 2009, when an unknown person (or group) with the pseudonym Satoshi Nakamoto released a whitepaper entitled: "Bitcoin: A Peer-to-Peer Electronic Cash System" [2]. The paper describes the possibility of creating a decentralised ledger in a peer to peer network (P2P) where, through the so-called "proof of work", the creation and exchange of Bitcoins would be allowed, which could be used as an innovative electronic currency. That same year, the P2P network began to operate and the first Bitcoins were created in an open source collaborative development environment. Nakamoto actively collaborated until mid 2010, when he/she/they transferred the source code repository and the project domains to the Bitcoin community and then disappeared as mysteriously as he/she/they appeared. This did not weaken the project, and gradually this idea became to be considered as one of the most disruptive in recent years with “The Economist” bringing the subject to worldwide attention with its cover story titled "The trust machine: the promise of the blockchain" and by awarding the technology its innovation prize in 2015 [3]. It is interesting to highlight this prize because, although the Bitcoin symbol appears in the front-page cover, the title refers to the technology behind it: Blockchain, which until then was mentioned only by experts.

Blockchain is the essential disruptive technology (rather than Bitcoin itself), since it allows the elimination of third parties in the transmission of money (such as central banks) and maintains the privacy of the user. Its innovative importance is so relevant that it has begun to be used in a series of applications that go beyond cryptocurrencies, there are even proposals to be implemented in the patent and trademark management systems themselves [4]. This versatility and the fact that it is an open source technology, has encouraged the development of many applications within the Bitcoin network, but also the creation of many other cryptocurrencies that compete / complement Bitcoin.

Today, hundreds of digital currencies based on Blockchain have been created with different features and aims and applications have been developed in numerous industries as a cost-effective and secure technique to create and manage a distributed database and maintain records for digital transactions of all types, since Blockchain-based transactions create quick, inexpensive and secure public records that can be used for many non-financial tasks, such as casting votes in elections or proving that a document existed at a specific time [5].

Given this enormous potential, many companies dealing with Blockchain technologies have emerged over the last years and a rush on patenting Blockchain technology began [6]. This makes this hot topic, a highly interesting field of study for a patent analysis in order to obtain insight into the developments of this technology.

Patent landscaping is a common tool to identify the evolution of a specific technical field or technologically hot, or emerging topic e.g. in the field of biotechnology [7] nanotechnology [8] [9] robotics [10] or additive manufacturing [11]. Prediction of emergent technologies within disruptive domains like Bitcoin has been done with a keyword network study analysing scholarly articles and business publications [12]. Patent landscaping studies of Blockchain exist, but have limitations such as either a small dataset [13], keyword-only search [14], a limited number of Blockchain related patent classifications [15] or only patents filed in a specific country e.g. the United States [16] [17]. We found only two studies which used a more complex and detailed search strategy to include multiple keywords, patent classifications and a

strategy to exclude false positives [18] [19]. Another interesting study used machine learning techniques for the classification of patents related with financial technologies [18], however, the paper does not go into sufficient detail in describing how the Blockchain patents are retrieved. Another interesting work qualifies Blockchain as GPT (General Purpose Technology), a key technology in the evolution of humanity of which there are only 24 considered so far in all of human history [20]. These authors infer that Blockchain is indeed a GPT through a longitudinal study of patents [22]. Unlike the previous work, which only studied financial technologies, this paper analyses Blockchain patents in all thematic fields and shows that although in 2014 more than 70% of patents were related to cryptocurrencies, this percentage fell to 40% in 2017 and it concludes that Blockchain is slowly becoming a GPT. The methodology of the data retrieval is in this case much more detailed than in the previous case and it combines keywords and IPC patent classification codes.

None of the hitherto mentioned studies use the number of keywords and classifications, and hence sophistication, described in this article and none describe a keyword weighting and a keyword/class validation by subject matter experts. In other words our study is based on a uniquely refined search strategy and resultant dataset.

2. Methodology

Search Strategy

In every bibliometric analysis one of the first steps is the definition of a search strategy in order to retrieve the most relevant set of documents. These documents, at a later stage, will form the data set to be analysed statistically. The identification of relevant keywords is a long and complex task, and there are different approaches to its solution.

In many cases bibliometric studies are carried out by researchers from a library/information science background, when subject matter experts who could suggest keywords beforehand are not available. Without subject matter experts, authors often build the set of keywords based on preliminary search results, creating a keyword starting set by iterative searches and a relevance feedback.

In our case, we were privileged to count on the knowledge of one of the most consolidated groups of subject matter experts in Europe: examiners from the European Patent Office (EPO). The EPO, with approximately 4,300 examiners, is one of the largest intellectual property offices in the world. This large and expert employee cohort allows each examiner to be highly specialised in a relatively small domain of technology. We collaborated with two examiners (see acknowledgments section) who contributed with valuable input and validated the keywords and patent classifications described in this article.

Blockchain patent classifications:

Patents are classified by their technological field, the most important classification schemes being the International Patent Classification (IPC) [1] and the Cooperative Patent Classification (CPC) [2]. CPC is a joint classification scheme of the European Patent Office and the United States' Patent and Trademark Office and is based on IPC but has significantly deeper levels of hierarchy and thus can define and describe the technology in patents in a more detailed way. Unfortunately not all patents are classified with the CPC whereas patent documents from most patent authorities worldwide are classified with the IPC.

[1] https://www.wipo.int/classifications/ipc/en/

[2] https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/classification/cpc.html

Currently, there is no IPC or CPC classification that clearly and unequivocally delimits the Blockchain thematic field. For this reason, the search strategy had to consist of a set of keywords and patent classifications combined in a way to retrieve the most relevant documents of the given search objective.

The following table (table 2) shows the classification codes of technologies with a collateral relation with Blockchain technologies, all of them belonging to the CPC which was shown to have more specific codes describing these technologies than the IPC. For every classification code we did a preliminary patent search in order to get an idea about the amount of patents that are using the class. The patent database Global Patent Index (GPI[3]) from the EPO with worldwide coverage was used for this purpose.

We found that in five cases, by combining two classes, the results were highly relevant (grey category in table 2) and no further refinement was needed. All other classes had to be combined with keywords in order to retrieve relevant results (blue category in table 1). Relevance was determined by expert, manual, qualitative assessment of the result sets.

Class	Description (Class)	How to use	Worldwide patent families	Scheme
H04L9/3247	Cryptographic mechanisms	in combination with keywords	5024	CPC
H04L9/3249	...involving digital signatures using RSA or related signature schemes, e.g. Rabin scheme	in combination with keywords	425	CPC
H04L9/3252	...involving digital signatures using DSA or related signature schemes, e.g. elliptic based signatures, ElGamal or Schnorr schemes	in combination with keywords	255	CPC
H04L9/3255	...involving digital signatures using group based signatures, e.g. ring or threshold signatures	in combination with keywords	198	CPC
H04L9/3257	...involving digital signatures using blind signatures	in combination with keywords	112	CPC
G06Q20/065	using e-cash	in combination with keywords	801	CPC
G06Q20/0652	...using e-cash...e-cash with decreasing value according to a parameter, e.g. time	in combination with keywords	140	CPC
G06Q20/0655	...using e-cash...e-cash managed centrally	in combination with keywords	301	CPC
G06Q20/0658	...using e-cash...e-cash managed locally	in combination with keywords	111	CPC
G06Q20/02	involving a neutral party	in combination with keywords	5301	IPC
G06Q20/023	...characterized in that the neutral party is a clearing house	in combination with keywords	612	CPC
G06Q20/4014	...Identity check for transaction	in combination with keywords	2595	CPC
G06Q20/4016	...involving fraud or risk level assessment in transaction processing	in combination with keywords	1759	CPC
G06Q20/4018	...using the card verification value [CVV] associated with the card	in combination with keywords	248	CPC
H04L9/3236	using cryptographic hash functions	in combination with H04L2209/56	1755	CPC
H04L9/3239	...involving non-keyed hash functions, e.g. modification detection codes [MDCs], MD5, SHA or RIPEMD	in combination with H04L2209/56	513	CPC
H04L9/3242	...involving keyed hash functions, e.g. message authentication codes [MACs], CBC-MAC or HMAC	in combination with H04L2209/56	1181	CPC
H04L9/0637	Modes of operation, e.g. cipher block chaining [CBC], electronic codebook [ECB] or Galois/counter mode [GCM]	in combination with keywords	683	CPC
H04L9/0643	Hash functions, e.g. MD5, SHA, HMAC or f9 MAC	in combination with keywords	1065	CPC
H04L2209/38	Chaining, e.g. hash chain or certificate chain	in combination with H04L2209/56	645	CPC
H04L2209/56	Financial cryptography, e.g. electronic payment or e-cash	in combination with H04L2209/38	2900	CPC
H04L2209/30	Compression, e.g. Merkle-Damgard construction	in combination with keywords	460	CPC
H04L2209/46	Secure multiparty computation, e.g. millionaire problem	in combination with keywords	134	CPC
H04L2209/463	...electronic voting	in combination with keywords	110	CPC
H04L2209/466	…electronic auction	in combination with keywords	19	CPC

Table 1: Classification analysis. Source: GPI

[3] https://www.epo.org/searching-for-patents/technical/espacenet/gpi.html

Blockchain keywords:
At first the keywords provided by the subject matter experts from the EPO were assigned to one of the 3 major concepts of Blockchain to which they belong: crypto-currencies, smart contracts and ledgers (table 2, first column).
Then they were analysed individually to check if they retrieved relevant Blockchain patents. For this purpose patents were searched in the database GPI that contained the specific keyword in either title or abstract (table 2 column 7) and subsequently the results were analysed and the keyword was given one of the following evaluations (in "result feedbacks" table 2 column 6):

- *High relevance*: meaning that most patents retrieved with this keyword were Blockchain relevant)
- *Relevant, but lots of noise:* meaning that the results contained relevant results but also many false positive (non relevant) patents
- *Too much noise* : meaning that most patents were not Blockchain relevant
- *No results:* meaning that the used keyword did not yield any patent documents

According to these results we then defined the following 3 categories of Blockchain keywords (table 2 column 2):

Main keywords (1):
These highly relevant keywords represent the core of our query. To filter out the few remaining false positives (e.g. patents describing a **chain**-saw for wooden **block**s) these keywords were crossed with one of the two very broad IPC/CPC classes that represent the ICT field: G06 (COMPUTING CALCULATING COUNTING) or H04 (ELECTRIC COMMUNICATION TECHNIQUES).

Secondary keywords (2):
These keywords are also relevant, but due to their semantic and contextual nuances they tend to generate many false positives. To avoid these we proposed combining them with the Blockchain relevant classes described in table 1 (grey category). In this way we were able to reduce the noise in a quite satisfactory way.

Discarded keywords (3):
These keywords, although associated with Blockchain, either retrieved no patents or very high percentage of non relevant patents for which even a combination with the before mentioned Blockchain relevant classes did not yield satisfactory results. For this reason they were excluded from the query.

Concept	Keyword Category	Keyword	Truncation/ Synonyms / Name Variations	How to use	Result feedback	Worldwide patent families (title & abstract search)
Crypto Currencies	1 Main keyword	BITCOIN	BITCOIN* OR BIT-COIN*	in combination with G06 OR H04	High relevance	167

Ledgers	1 Main keyword	BLOCKCHAIN	BLOCKCHAIN* OR BLOCK-CHAIN*	in combination with G06 OR H04	High relevance	5053
Ledgers	1 Main keyword	BLOCKSIGN	BLOCKSIGN	in combination with G06 OR H04	High relevance	2
Smart Contracts	1 Main keyword	CODIUS	CODIUS	in combination with G06 OR H04	High relevance	1
Crypto Currencies	1 Main keyword	COLORED COINS	COLORED-COIN* OR COLOURED-COIN*	in combination with G06 OR H04	High relevance	60
Crypto Currencies	1 Main keyword	CRYPTO CURRENCY	CRYPTOCURRENC* OR CRYPTO-CURRENC*	in combination with G06 OR H04	High relevance	819
Ledgers	1 Main keyword	DISTRIBUTED LEDGER	DISTRIBUTED-LEDGER	in combination with G06 OR H04	High relevance	376
Crypto Currencies	1 Main keyword	DOGECOIN	DOGECOIN OR DOGE-COIN	in combination with G06 OR H04	High relevance	54
Crypto Currencies	1 Main keyword	ETHEREUM	ETHEREUM	in combination with G06 OR H04	High relevance	214
Ledgers	1 Main keyword	FACTOM	FACTOM	in combination with G06 OR H04	High relevance	10
Crypto Currencies	1 Main keyword	LITECOIN	LITECOIN OR LITE-COIN	in combination with G06 OR H04	High relevance	180
Ledgers	1 Main keyword	P2SH	PAY-TO-SCRIPT-HASH OR P2SH	in combination with G06 OR H04	High relevance	46
Ledgers	1 Main keyword	PROOF OF STAKE	PROOF-OF-STAKE	in combination with G06 OR H04	High relevance	154
Ledgers	1 Main keyword	SIDECHAIN	SIDECHAIN*	in combination with G06 OR H04	High relevance	67
Smart Contracts	1 Main keyword	SMART CONTRACT	SMART-CONTRACT* OR SMARTCONTRACT*	in combination with G06 OR H04	High relevance	583
Crypto Currencies	1 Main keyword	ZEROCASH	ZEROCASH OR ZCASH	in combination with G06 OR H04	High relevance	14
Smart Contracts	2 Secondary keyword	CHAINCODE	CHAINCOD*	in combination with classes	Relevant, but lots of noise	9
Crypto Currencies	2 Secondary keyword	COUNTERPARTY	COUNTERPARTY OR XCP	in combination with classes	Relevant, but lots of noise	3238
Crypto Currencies	2 Secondary keyword	DIGITAL CURRENCY	DIGITALCURRENC* OR DIGITAL-CURRENC*	in combination with classes	Relevant, but lots of noise	1187
Crypto Currencies	2 Secondary keyword	ETHER	ETHER	in combination with classes	Relevant, but lots of noise	6898
Ledgers	2 Secondary keyword	FORKING	FORKING OR FORKS	in combination with classes	Relevant, but lots of noise	2541
Smart Contracts	2 Secondary keyword	HAWK	HAWK	in combination with classes	Relevant, but lots of noise	620
Ledgers	2 Secondary keyword	LEDGER	LEDGER*	in combination with classes	Relevant, but lots of noise	10413
Crypto Currencies	2 Secondary keyword	LISK	LISK	in combination with classes	Relevant, but lots of noise	20
Ledgers	2 Secondary keyword	MERKLE TREE	MERKLE-TREE OR MERKLETREE OR HASH-TREE OR HASHTREE OR MERKLE-ROOT OR MERKLEROOT	in combination with classes	Relevant, but lots of noise	868
Crypto Currencies	2 Secondary keyword	METACOIN	META-COIN* OR METACOIN*	in combination with classes	Relevant, but lots of noise	3
Crypto Currencies	2 Secondary keyword	NAMECOIN	NAME-COIN* OR NAMECOIN*	in combination with classes	Relevant, but lots of noise	529
Crypto Currencies	2 Secondary keyword	NXT	NXT	in combination with classes	Relevant, but lots of noise	841
Ledgers	2 Secondary keyword	PROOF OF WORK	PROOF-OF-WORK OR HASH-CASH OR HASHCASH	in combination with classes	Relevant, but lots of noise	581
Ledgers	2 Secondary keyword	RIPPLE	RIPPLE	in combination with classes	Relevant, but lots of noise	16192
Smart Contracts	2 Secondary keyword	ROOTSTOCK	ROOTSTOCK OR RSK	in combination with classes	Relevant, but lots of noise	228
Crypto Currencies	2 Secondary keyword	STELLAR	STELLAR	in combination with classes	Relevant, but lots of noise	221
Smart Contracts	2 Secondary keyword	SYMBIONT	SYMBIONT	in combination with classes	Relevant, but lots of noise	18
Crypto Currencies	2 Secondary keyword	TYPECOIN	TYPE-COIN* OR TYPECOIN*	in combination with classes	Relevant, but lots of noise	313

Crypto Currencies	2 Secondary keyword	ZEROCOIN	ZEROCOIN OR ZERO-COIN	in combination with classes	Relevant, but lots of noise	9
Ledgers	2 Secondary keyword	ZEROKNOWLEDGE	ZEROKNOWLEDGE OR ZERO-KNOWLEDGE	in combination with classes	Relevant, but lots of noise	1592
Crypto Currencies	3 Discarded keyword	BAMBOO	BAMBOO	not to use	Too much noise	1622
Crypto Currencies	3 Discarded keyword	CARDANO	CARDANO	not to use	Too much noise	13
Crypto Currencies	3 Discarded keyword	GAS	GAS OR GAS-PRICE	not to use	Too much noise	65469
Crypto Currencies	3 Discarded keyword	BITHOMP	BITHOMP	not to use	No results	0
Smart Contracts	3 Discarded keyword	CLICKTOPURCHASE	CLICKTOPURCHASE	not to use	No results	0
Crypto Currencies	3 Discarded keyword	DOGEPARTY	DOGEPARTY	not to use	No results	0
Smart Contracts	3 Discarded keyword	MONAX	MONAX	not to use	No results	0
Smart Contracts	3 Discarded keyword	TEZOS	TEZOS	not to use	No results	0
Ledgers	3 Discarded keyword	HASH FUNCTION	HASH-FUNCTION	not to use	Too much noise	29061
Ledgers	3 Discarded keyword	MINING	MINING OR MINER*	not to use	Too much noise	39033
Ledgers	3 Discarded keyword	SHA/SHA256	SECURE-HASH-ALGORITHM OR SHA OR SHA256 OR SHA-256	not to use	Too much noise	17476

Table 2: Keyword analysis. Source: GPI

Figure 1 below shows the top 28 keywords and their results of the preliminary search. The colours indicate the keyword category. It becomes obvious that the discarded keywords of category 3 (in red) introduce a lot of noise, especially the first four. They are followed by three keywords of category 2 (in yellow) that are secondary keywords and whose noise can be satisfactory reduced by combining them with classifications. The first main keyword that is obviously Blockchain appears only in eighth place. In lower positions in the chart are all main or secondary keywords, but it can be seen that these keywords retrieve only a small number of patents.

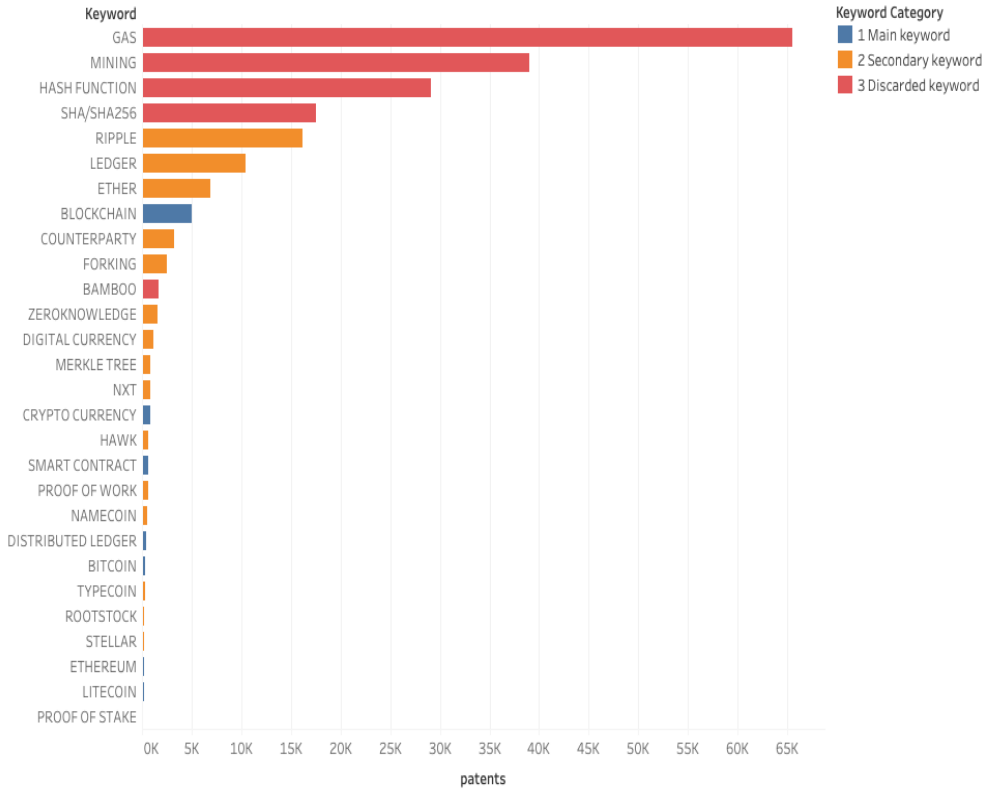

Fig. 1: Keywords preliminary search by type. Source: GPI

Blockchain query proposal:

Using the relevant classifications and keywords as described previously we proposed the following query to be used for a search of Blockchain related patents, and with 4 sub-queries as follows:

- Sub-Query 1: this query combines the main keywords (keyword category 1) with the two very broad IPC/CPC classes that represent the ICT field: G06 or H04
- Sub-Query 2: this query combines the secondary keywords (keyword category 2) with the Blockchain related classifications (blue category, table 1)
- Sub-Query 3: this query combines the Blockchain related classifications, whose results were highly relevant and no further refinement with keywords was needed (grey category in table 2)
- Full Query: this is the final query and is a combination of the sub-queries 1 to 3 with a Boolean OR logic

The query proposal is using the Query syntax of GPI but can be adapted to every patent database that allows multiple keyword and class search.

Sub-Query 1 (Q1):
WORD=(BLOCKCHAIN* OR "BLOCK-CHAIN" OR BITCOIN* OR "BIT-COIN" OR BLOCKSIGN OR CODIUS OR "COLORED-COIN" OR "COLOURED-COIN" OR CRYPTOCURRENC* OR "CRYPTO-CURRENC" OR "DISTRIBUTED LEDGER" OR DOGECOIN OR "DOGE-COIN" OR ETHEREUM OR FACTOM OR LITECOIN OR "LITE-COIN" OR "PAY-TO-SCRIPT-HASH" OR P2SH OR "PROOF-OF-STAKE" OR SIDECHAIN* OR "SMART-CONTRACT" OR SMARTCONTRACT* OR ZEROCASH OR ZCASH) AND CLAS=(G06 OR H04)

Sub-Query 2 (Q2):
WORD=(CHAINCOD* OR COUNTERPARTY OR XCP OR DIGITALCURRENC* OR "DIGITAL-CURRENC" OR ETHER OR FORKING OR FORKS OR HAWK OR LEDGER OR LISK OR "MERKLE-TREE" OR MERKLETREE OR "HASH-TREE" OR HASHTREE OR "MERKLE-ROOT" OR MERKLEROOT OR "META-COIN" OR METACOIN* OR "NAME-COIN" OR NAMECOIN* OR NXT OR "PROOF-OF-WORK" OR "HASH-CASH" OR HASHCASH OR ROOTSTOCK OR RSK OR RIPPLE OR STELLAR OR SYMBIONT OR "TYPE-COIN" OR TYPECOIN* OR ZEROCOIN OR "ZERO-COIN" OR ZEROKNOWLEDGE OR "ZERO-KNOWLEDGE") AND CLAS=(H04L9/3247 OR H04L9/3249 OR H04L9/3252 OR H04L9/3255 OR H04L9/3257 OR H04L9/3236 OR H04L9/3239 OR H04L9/3242 OR H04L9/0637 OR H04L9/0643 OR H04L2209/38 OR H04L2209/56 OR H04L2209/30 OR H04L2209/46 OR H04L2209/463 OR H04L2209/466 OR G06Q20/065 OR G06Q20/0652 OR G06Q20/0655 OR G06Q20/0658 OR G06Q20/02 OR G06Q20/023 OR G06Q20/027 OR G06Q20/401 OR G06Q20/4012 OR G06Q20/4014 OR G06Q20/40145 OR G06Q20/4016 OR G06Q20/4018)

Sub-Query 3 (Q3):
CLAS=((H04L9/3236 OR H04L9/3239 OR H04L9/3242 OR H04L2209/38) AND H04L2209/56)

Full Query ((Q1 OR Q2 OR Q3):
(WORD=(BLOCKCHAIN* OR "BLOCK-CHAIN" OR BITCOIN* OR "BIT-COIN" OR BLOCKSIGN OR CODIUS OR "COLORED-COIN" OR "COLOURED-COIN" OR CRYPTOCURRENC* OR "CRYPTO-CURRENC" OR "DISTRIBUTED LEDGER" OR DOGECOIN OR "DOGE-COIN" OR ETHEREUM OR FACTOM OR LITECOIN OR "LITE-COIN" OR "PAY-TO-SCRIPT-HASH" OR P2SH OR "PROOF-OF-STAKE" OR SIDECHAIN* OR "SMART-CONTRACT" OR SMARTCONTRACT* OR ZEROCASH OR ZCASH) AND CLAS=(G06 OR H04) OR WORD=(CHAINCOD* OR COUNTERPARTY OR XCP OR DIGITALCURRENC* OR "DIGITAL-CURRENC" OR ETHER OR FORKING OR FORKS OR HAWK OR LEDGER OR LISK OR "MERKLE-TREE" OR MERKLETREE OR "HASH-TREE" OR HASHTREE OR "MERKLE-ROOT" OR MERKLEROOT OR "META-COIN" OR METACOIN* OR "NAME-COIN" OR NAMECOIN* OR NXT OR "PROOF-OF-WORK" OR "HASH-CASH" OR HASHCASH OR ROOTSTOCK OR RSK OR RIPPLE OR STELLAR OR SYMBIONT OR "TYPE-COIN" OR TYPECOIN* OR ZEROCOIN OR "ZERO-COIN" OR ZEROKNOWLEDGE OR "ZERO-KNOWLEDGE") AND CLAS=(H04L9/3247 OR H04L9/3249 OR H04L9/3252 OR H04L9/3255 OR H04L9/3257 OR H04L9/3236 OR H04L9/3239 OR H04L9/3242 OR H04L9/0637 OR H04L9/0643 OR H04L2209/38 OR H04L2209/56 OR H04L2209/30 OR H04L2209/46 OR H04L2209/463 OR H04L2209/466 OR G06Q20/065 OR G06Q20/0652 OR G06Q20/0655 OR G06Q20/0658 OR G06Q20/02 OR G06Q20/023 OR G06Q20/027 OR G06Q20/401 OR G06Q20/4012 OR G06Q20/4014 OR G06Q20/40145 OR G06Q20/4016 OR G06Q20/4018) OR CLAS=((H04L9/3236 OR H04L9/3239 OR H04L9/3242 OR H04L2209/38) AND H04L2209/56))

Table 3: Blockchain sub-queries Q1-Q4 and full query in GPI query format

Case study: worldwide Blockchain patent landscape 2008-2018

In order to test our query we carried out a patent landscaping exercise with methodology as described as follows:

- GPI from EPO[4] and Orbit Intelligence from Questel[5] were used for data retrieval and analysis. Charts, rankings & diagrams were created with Microsoft Excel. Citation node & collaboration maps were created with Orbit Intelligence.
- Patent publications were counted from the year 2008 on, since this was the year when the digital currency Bitcoin was disclosed via a whitepaper and to date Bitcoin is the most successful implementation of Blockchain technology.
- Patents were counted as simple patent families (one count per invention) and full count was applied.
- The patent applicants were considered for the top player analysis and citation node map. As the analysed patent data originated from multiple patent authorities worldwide, misspellings and different name variations occur. This made it necessary to manually correct the inventor and applicant names. Furthermore company names were standardised based on known mergers, acquisitions and change of ownership and sub-companies of big corporations were aggregated to the main company name. This time

[4] https://www.epo.org/searching-for-patents/technical/espacenet/gpi.html
[5] https://www.questel.com/ip-business-intelligence-software/orbit-intelligence/

consuming and tedious task was completed with the help of the "data-rules" tool from Orbit that helped to find similar names and spellings via automated semantic analysis of the applicant names of the data set.

This analysis was done for exemplary purposes only in order to demonstrate the developed search methodology. Its purpose is not to be a comprehensive patent landscape analysis, but to show how the search strategy can be applied and to demonstrate how the analysis of patent information can be used to gain insight into a specific complex technological field like Blockchain.

3. Results of case study

Our research, which was completed in October 2018, generated a data set of 4096 patent families. Selected results have been presented at the EPO's Patenting Blockchain conference [23], at the EPO's seminar "Search Matters" [24] and at the Patent Information Users' Group annual conference [25]. A subsequent cursory analysis in February 2019 of our data, using our search approach indicated that the number of families had approximately doubled in 3-4 months [26].

Overview and evolution

Blockchain patenting took off in 2015, since then we observe high publication rates, and steady growth of granted patents (Fig. 2).

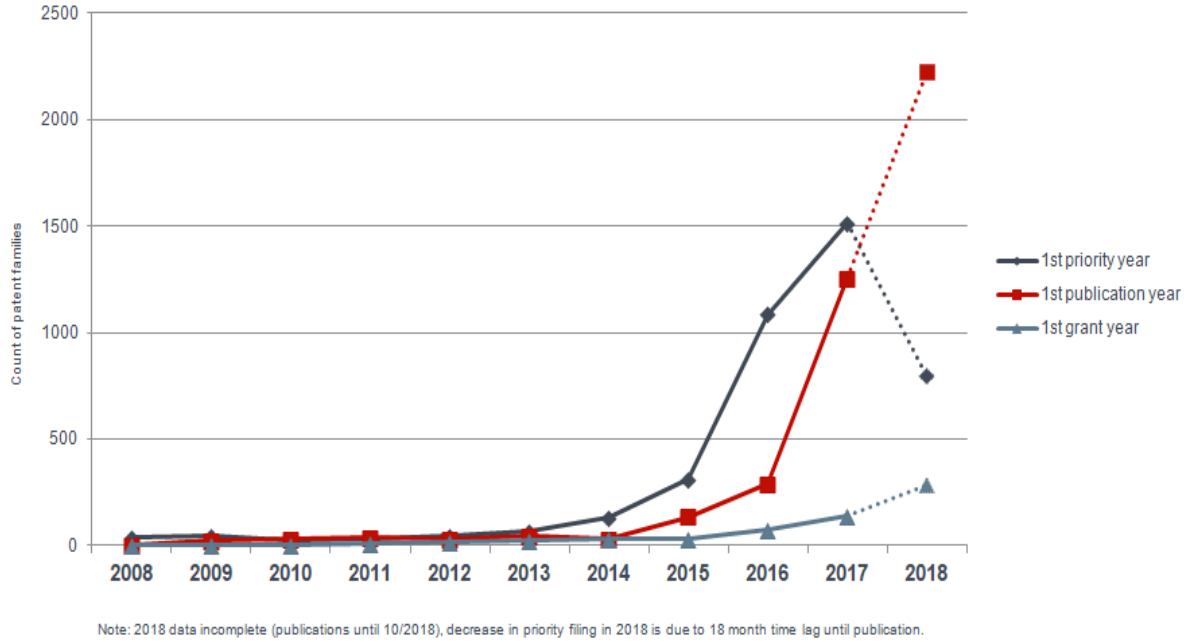

Fig. 2: Blockchain patenting evolution[6]

The high growth of patent applications at the CNIPA (China) authority since 2016 and steadily increasing rate of Blockchain patent application in USPTO (US), WIPO (WO[7]) and KIPO

[6] 1 st priority year: earliest year when a patent application for the invention was filed (priority patent) /1st publication year: earliest year when a patent of the invention was published / 1st grant year: earliest year when a patent of the invention was granted

[7] WO corresponds to PCT patent applications. Managed by the World Intellectual Property Organization (WIPO) the PCT route is often used by multinational companies and research organisations since it offers the possibility to extend a patent to multiple countries worldwide with a single application. It is called PCT application and not a PCT patent since the granting of patents remains under the control of the national patent offices.

(South Korea) are shown in Figure 3. Here we see clearly that the Chinese office publications are not only ahead in absolute values (around 40% of the applications), but also present a divergent behaviour, with a growth curve much greater than the rest. The US office is second with around 20% of all the applications.

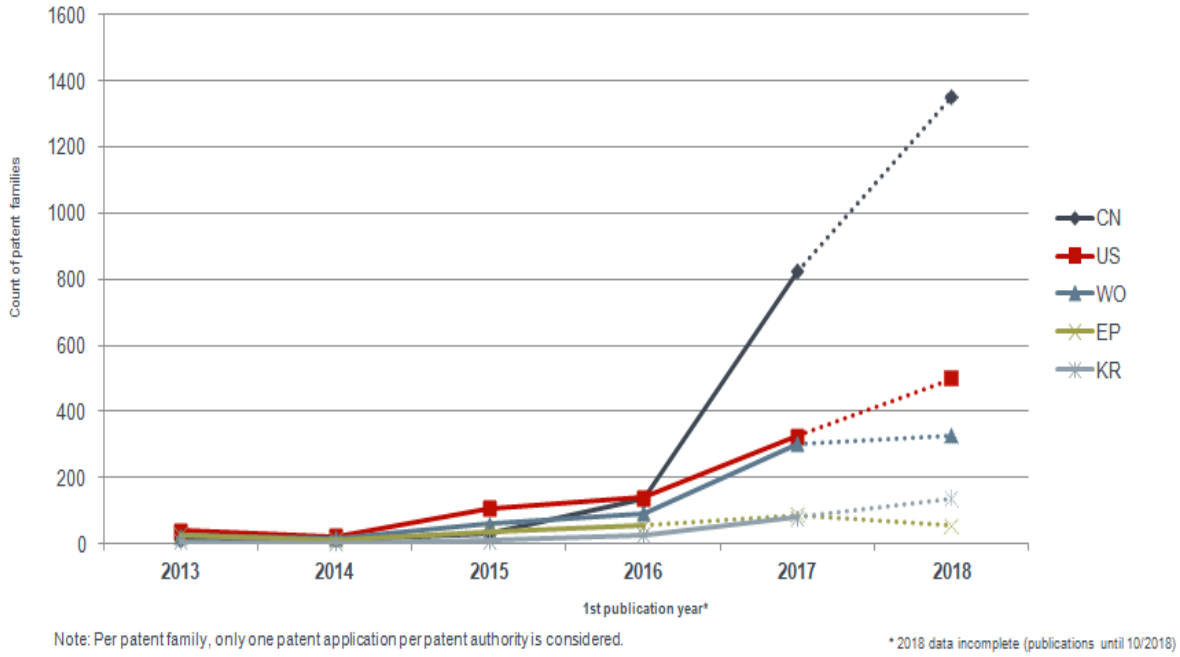

Fig. 3: Patent applications per patent authority (2013-2018)

We then compared the country results in a relative manner by calculating their "Specialisation ratio" (SR). We define this ratio as the number of priority patent applications of a country in a specific technological field (here: Blockchain) p(t) in a period t, divided by the total number of priority applications of all priority patent applications of that country P(t) in the same period:

$SR = p(t)/P(t)$

In our opinion this ratio can be used to indicate the degree of importance that a particular technology has in a particular country or jurisdiction since compares the patent applications related to a certain technology field to the total applications that a specific office / jurisdiction receives and thus a bias can indicate an important or disputed market for the technology.

In the context of Blockchain we found that three Anglo-Saxon countries have a higher specialisation ratio in Blockchain patent applications (Table 4).

Patent Authority	**Blockchain patent applications (published 2008-2018)**	**Total patent applications (published 2008-2018)**	**Specialisation ratio**
IPAustralia	45	69791	0,64 ‰
USPTO	1089	2375381	0,45 ‰
UKIPO	104	239630	0,43 ‰
WIPO	104	244348	0,42 ‰
EPO	73	251160	0,29 ‰
CNIPA	2198	13185169	0,16 ‰
KIPO	221	1558994	0,14 ‰

JPO	72	3548358	0,02 ‰

Table 4 – Ratio of Blockchain specialisation by patent authority

The 45 applications of Australia (IPAustralia) may seem small but they are 0.64 ‰ of all patent applications in this country, with Australia thus leading the specialisation ranking. A little behind we found United States (USPTO) and United Kingdom (UKIPO), each with more than 0.4 ‰. Subsequently come the two multinational offices (WIPO and EPO), and far behind, the Asian offices. In this sense, the very low degree of specialisation of Japan (JPO) is significant, despite being a world giant in the field of patents, Japan barely appears in the Blockchain patent world.

Table 5 below shows a top 20 list of the applicants with most Blockchain related applications. All of them are private companies, 12 are Chinese, and half of them are listed on the stock exchange, the majority above the $100 billion of market capitalisation. In the case of the US, the five companies are widely known in the world of finance (Bank of America, MasterCard and VISA) or technology (IBM and Intel). Altogether among the five they add more than 1300 billions in market capitalization. The first three Chinese applicants are large companies in the stock market (Alibaba, BOE and Unicom) with a total of more than $900 billions, however, the remaining nine are much smaller technological companies. In the case of Korea and the UK they also appear to be small tech companies. We only identified one company from continental Europe, the Finnish Nokia (acquired by Microsoft).

Applicant	**Country of residence**	**Patent families**
IBM	US	111
ALIBABA HOLDING	CN	88
COINPLUG	KR	88
BOE TECHNOLOGY	CN	61
MASTERCARD	US	51
BANK OF AMERICA	US	46
CHINA UNICOM	CN	46
NCHAIN HOLDINGS	UK	45
VISA	US	41
HANGZHOU FUZAMEI TECHNOLOGY	CN	36
BEIJING RUI ZHUO XITONG TECHNOLOGY DEVELOPMENT	CN	34
JINAN WAVE OF HIGH TECH INVESTMENT & DEVELOPMENT	CN	32
INTEL	US	31
SINOCHAIN TECHNOLOGY	CN	30
TENCENT TECHNOLOGY	CN	29
HANGZHOU YUNPHANT NETWORK TECHNOLOGY	CN	28
BEIJING EUROPE CHAIN TECHNOLOGY	CN	27
CLOUDMINDS	CN	27
JIANGSU TONGFUDUN TECHNOLOGY	CN	27
NOKIA	FI	27

Table 5 – Top 20 applicants

Figure 4 shows a graph of the temporal evolution of patent applications of the main companies. It can be appreciated that Chinese Alibaba was the leader until 2017 although it slowed down recently. On the other hand the US-American IT giant IBM and the South Korean Bitcoin brokerage and service provider Coinplug appear as the companies with the biggest momentum with a considerable growth in Blockchain patent publications in the last years. The rest of the companies show a more balanced behaviour.

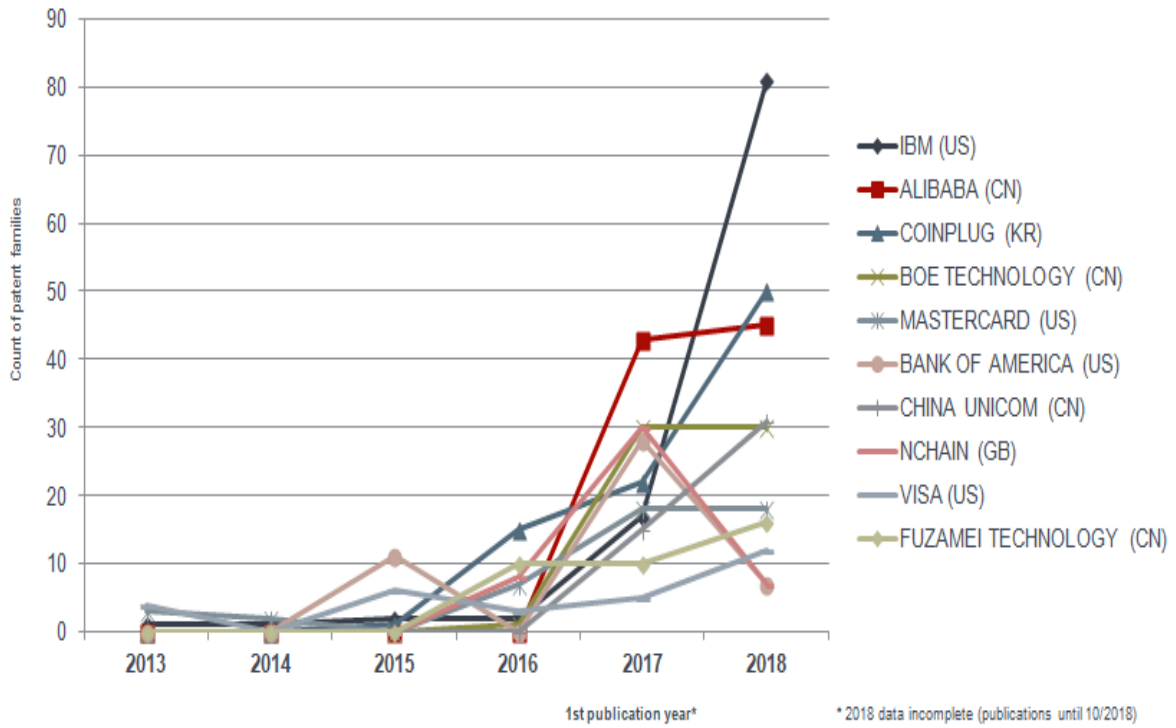

Fig. 4: Patent applications of top 10 applicants (2013-2018)

Another way of presenting the relationship between companies is through citation graphs. Similar to references in scientific publications, patent citations can reveal other documents that influenced in some way the invention, either by describing the state of the art (usually references from the inventor) but also documents that describe prior art (references identified by the patent examiner). We analysed the 4096 patent families of the dataset and visualised companies that have at least 20 patents and which have citations to/from another patents (Figure 5). The bubble size of a company corresponds to the Blockchain portfolio size (number in green circle) and the arrow with number indicates how many patents cite the applicant where the arrow is directed at.

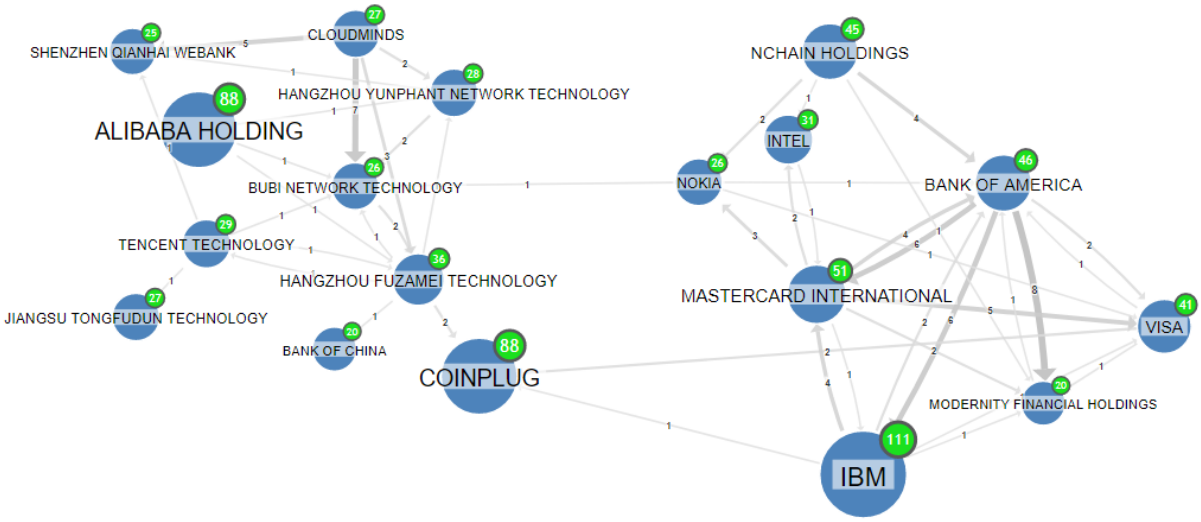

Fig. 5: Citation network node map. Source: Orbit

On the left an “Asian-centric” cluster becomes evident, with many patents of Chinese companies citing each other and with Bubi Network, a Chinese company that focuses on innovation of Blockchain related technologies, being the most cited one (12 citations received from 5 different applicants). For most other applicants in this cluster the intensity of the citation is low, with only one or two citations per link. BOE Technology does not appear since, despite a large portfolio, no citations were detected.

On the right side we can speak of an “Anglo-American centric” cluster with mainly US financial companies citing each other. Here in general the citations are more intense and highlight the centrality of the Blockchain portfolio of the Bank of America (also with 12 citations received from 5 different applicants) and whose patents cite those of several other related companies like Visa or Mastercard, but also the company Modernity Financial Holdings. This Taiwanese company, that operates a Bitcoin exchange service platform, does not appear in the ranking in table 5 (because it is in position 34), but it seems to have exercised a lot of influence with its patent portfolio.

Finally, it should be noted that both clusters are connected by two companies that can be considered as "hubs": Coinplug (being the only applicant cited by both, applicants from the Asian and Anglo-American cluster) and the European technology company Nokia (that cites an applicant from the Asian cluster).

4. Conclusion

Blockchain, Bitcoin and all related technologies are in a process of high growth and innovation. Some authors even call this field of knowledge a “General Purpose Technology” (GPT), that is, a key technology for the evolution of Humanity. Naturally such an important technological field has a great impact in terms of patent applications. However, there is no patent classification that clearly and unequivocally delimits the Blockchain thematic field. In this paper we have proposed the most complete query that has been published so far in the scientific literature for identifying Blockchain related patents.

To construct this query, we collected the opinion of subject matter experts and patent examiners from the European Patent Office. They proposed a group of keywords that we analysed against the patent database in order to establish those that were the main (major), the secondary and those that were eliminated (discarded). The criterion used here was the amount of false positives (noise) they generated, after combining them with the classification codes also proposed by the examiners.

To test our query we carried out a case study analysing worldwide Blockchain patents from 2008 to 2018. We found increasing patent growth from 2015 on, corresponding to the same dates that Bitcoin and Blockchain appeared on the cover of The Economist. From that time, we detect worldwide growth, although the office that receives the most applications, in absolute terms, is the CNPIA (China), with almost half of the total. However, if we measure it in relative terms (as a percentage of the total patent applications) we found that there are three Anglo-

Saxon offices that have a greater "Blockchain specialisation": Australia, United States and United Kingdom (in that order). The European specialisation rate (EPO) is half of the previous ones, although far ahead of China. We then analysed the top companies for patent applications, and found that the ones with large stock market capitalization stand out, such as the Americans of the world financed (Bank of America, Mastercard and VISA) and those of technology (IBM and Intel). In the case of Chinese companies there is also some large cap (as in the case of Alibaba), although many smaller start-ups appear in the field of financial technologies. The citation map revealed two differentiated geographical domains, which may be indicating two different ways of innovation. Finally, it should be noted that the European presence is limited to very few companies, so we could say that the Blockchain as a field of innovation has not yet matured in Europe.

References

1. Swan, M. (2015). Blockchain: blueprint for a new economy. Sebastopol. O'Reilly.
2. Nakamoto S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf Accessed: 19.03.2019
3. Economist (2015). The TRUST machine: the promise of the blockchain. The Economist, Oct 31st 2015. https://www.economist.com/leaders/2015/10/31/the-trust-machine Accessed: 15.03.2019
4. Alnafrah I., Bogdanova E.; Maximova T. (2019). Text mining as a facilitating tool for deploying blockchain technology in the intellectual property rights system. International Journal of Intellectual Property Management (IJIPM), Vol. 9, No. 2
5. Boucher P, et al (2017). How blockchain technology could change our lives, European Parliament Publication Service, Scientific Foresight Unit http://www.europarl.europa.eu/RegData/etudes/IDAN/2017/581948/EPRS_IDA(2017)581948_EN.pdf Accessed: 15.04.2019
6. Reuters (2017). Bitcoin's "creator" races to patent technology with gambling tycoon, In: https://www.reuters.com/investigates/special-report/bitcoin-wright-patents/ Accessed: 15.03.2019
7. Jürgens, B. and Clarke, N. (2019) Evolution of CAR T-cell immunotherapy in terms of patenting activity. Nature Biotechnology 37 (4), 370
8. Jürgens, B. & Herrero-Solana V. (2017a). Monitoring nanotechnology using patent classifications: an overview and comparison of nanotechnology classification schemes. Journal Nanopart Res 19: 151.
9. Jürgens, B. and Herrero-Solana, V. (2017b) Patent bibliometrics and its use for technology watch. Journal of Intelligence Studies in Business. 7 (2) 17-26.
10. Göldner, M., C. Herstatt and F. Tietze (2015). The emergence of care robotics – A patent and publication analysis. Technological Forecasting and Social Change 92(March): 115-131.
11. List, J. and F. Tietze (2017). Additive Manufacturing Innovations by UK Organisations: A review of patents and academic literature published between 2006 and 2015. Cambridge, UK, Institute for Manufacturing, University of Cambridge.

12. Dotsika, F., & Watkins, A. (2017). Identifying potentially disruptive trends by means of keyword network analysis. Technological Forecasting and Social Change, 119, 114-127.

13. Khurana T. (2017). Sample patent landscape study-blockchain , IIPRD, http://www.iiprd.com/wp-content/uploads/2015/11/IIPRD-Patent-Landscape-Study-Blockchain.pdf Accessed: 10.03.2019

14. Chiu H. (2017). An Overview of the Blockchain Patent Landscape, Clarivative Analytics, https://clarivate.com/derwent/blog/blockchain-disruption-or-hype-key-webinar-takeaways/ Accessed: 15.03.2020

15. Okutan E (2018) Blockchain Patent Classes that Deserve Attention, Patenteffect, 2018, https://www.patenteffect.com/post/blockchain-patents Accessed: 15.03.2020

16. Marshall (2017). The Emerging Blockchain Patent Landscape, Marshall IP, https://www.lexology.com/library/detail.aspx?g=cf0c71c5-055a-4d57-92f8-c75d1e282414 Accessed: 15.03.2020

17. Lee A (2018). Blockchain patent filings dominated by financial services industry, Patentvue, http://patentvue.com/2018/01/12/blockchain-patent-filings-dominated-by-financial-services-industry/ Accessed: 18.03.2019

18. Patsnap (2018). An IP perspective on Blockchain: beyond the realm of cryptocurrencies, Patsnap, https://info.patsnap.com/hubfs/Academy/Course Material/Technology-Landscape-Blockchain-Example.pdf Accessed: 15.03.2020

19. Patseer (2017). Using Patseer Pro to Search & Analyze Patents on Blockchain, Patseer, Gridlogics, https://patseer.com/2017/03/patent-landscape-report-on-blockchain-by-patseer-pro/ Accessed: 15.03.2019

20. Chen, M. A., Wu, Q., & Yang, B. (2019). How Valuable Is FinTech Innovation?. The Review of Financial Studies, 32(5), 2062-2106.

21. Lipsey, R. G., Carlaw, K. I., & Bekar, C. T. (2005). Economic transformations: general purpose technologies and long-term economic growth. Oxford University Press, pp. 131–218.

22. Filippova E., Scharl A., Filippov P. (2019) Blockchain: An Empirical Investigation of Its Scope for Improvement. In: Joshi J., Nepal S., Zhang Q., Zhang LJ. (eds) Blockchain – ICBC 2019. ICBC 2019. Lecture Notes in Computer Science, vol 11521. Springer, Cham. https://doi.org/10.1007/978-3-030-23404-1_1

23. Weber G. Menière Y (2018) EPO Conference Patenting Blockchain https://www.youtube.com/watch?v=ROwzOLxngsQ&list=PLgDlswd0__eLf-pkgullgnILJSjEjt3Sr&index=3&t=0s

24. Boedt G. EPO Search Matters 2019. Lecture 3. Patent landscaping methodology, a case study on Blockchain inventions https://e-courses.epo.org/pluginfile.php/42269/mod_resource/content/2/searchmatters2019/index.html

25. Clarke N. Patent Information Users' Group Annual Conference May 2019 "Patent Insight Reports from the EPO: Future and Emerging Technologies" https://www.piug.org/resources/Documents/an/2019/Workshop_abstract_EPO_AN2019.pdf

26. Clarke N., Kanesarajah V. White E. Clarivate Webinar Blockchain. Disruption or Hype? 15th May 2019 https://clarivate.com/derwent/campaigns/blockchain-disruptive-technology-or-just-a-flash-in-the-ip-pan/

Supplementary information

The opinions expressed in this article are those of the authors and not necessarily those of their employers.

Acknowledgements

We thank Thierry Bec and Marc Diepstraten from the European Patent Office for their expert support in the search strategy. The present article is based on a patent data study that was commissioned by the European Patent Office, Department 5.4.2.1 Patent Information research, with the aim of demonstrating to readers outside the disciplines of patent analysis and Blockchain technologies, how patent information can be used to gain insight and competitive advantages about a specific technological field. Additional bibliometric analysis, for the purposes of the present article, was supported by the Spanish Ministry of Science Innovation and Universities under the framework R&D project "Detection of Spanish technology for connected Industry 4.0 in European patent offices" (RTI2018-098966-B-I00).

Competing Interests

The authors declare no competing financial interests