Global innovation and competition in quantum technology, viewed through the lens of patents and artificial intelligence

Zeki Can Seskir

Institute for Technology Assessment and Systems Analysis,
Karlsruhe Institute of Technology,
Karlstraße 11, 76133 Karlsruhe, Germany
Email: zeki.seskir@kit.edu

Kelvin W. Willoughby*

Deutsche Bank Chair of Innovation Management and Entrepreneurship,
HHL Leipzig Graduate School of Management,
Jahnallee 59, 04109 Leipzig, Germany
Email: k.willoughby@hhl.de
*Corresponding author

Abstract: In this work we elucidate international trends in the field of quantum technology (QT) by analysing a global patent database built from an operational definition of QT that was generated through the curated application of artificial intelligence (AI). In doing so, we demonstrate how the sophisticated use of intellectual property information, enhanced by the artful deployment of AI techniques, may produce more reliable and useful revelations for policymakers and managers about global innovation in emerging fields of technology than is possible through conventional methods of data collection and analysis. We also demonstrate the utility of this approach for reliably characterising the evolving constituent sub-fields of QT. By adopting a hybrid human-AI approach to both the definition and the analysis of QT, we have produced some novel insights about global innovation and national organisational profiles in the QT field, particularly concerning dynamic competition between the USA and China.

Keywords: quantum technology; quantum innovation; patent analysis; artificial intelligence; patent landscape; patinformatics; global technological innovation; quantum competition; quantum industry.

Reference to this paper should be made as follows: Seskir, Z.C. and Willoughby, K.W. (2023) 'Global innovation and competition in quantum technology, viewed through the lens of patents and artificial intelligence', *Int. J. Intellectual Property Management*, Vol. 13, No. 1, pp.40–61.

Biographical notes: Zeki Can Seskir has education and degrees in both Physics and Science and Technology Policy Studies. He is currently a PhD candidate at the Karlsruhe Institute of Technology, Institute for Technology Assessment and Systems Analysis, working on technology and vision assessment for quantum technologies. He is also a board member of QWorld, an independent quantum non-government-organisation active in over twenty countries. His research interests cover technology assessment, science and technology studies, and the ethical, legal and societal aspects of research on quantum technologies.

Kelvin W. Willoughby holds the Deutsche Bank Chair of Innovation Management and Entrepreneurship at the HHL Leipzig Graduate School of Management. He is an expert on the management of intellectual property, technology-based entrepreneurship, and strategic planning for technology-based industry development. He holds Doctorates in both Strategic Management and Technology studies, and a Master of Laws degree in Intellectual Property Law. He has extensive experience as an educator, researcher and industry consultant in the USA, Europe, Asia, Australia and Russia. He is also an IEEE Senior Member.

This paper is a revised and expanded version of a paper entitled 'An approach to AI-enhanced patent analysis for emerging advanced technology industries' presented at the 81st Annual Meeting of the Academy of Management, virtual meeting, 30 July–3 August 2021.

1 Introduction

Quantum technology (QT) is a field of innovation attracting global attention in recent years. Two initiatives, each in the order of one billion Euro (€1B), were recently enacted by the EU (Riedel et al., 2019) and the USA (Raymer and Monroe, 2019). Similar programs exist, all initiated within the last decade, in Canada (Sussman et al., 2019), Japan (Yamamoto et al., 2019), Australia (Roberson and White, 2019), the UK (Knight and Walmsley, 2019), Russia (Fedorov et al., 2019), and China (Zhang et al., 2019). The combined planned investment through such initiatives worldwide is currently around €20B (Qureca, 2020). One of the main motivations for these initiatives is to support the commercialisation of quantum technologies, an activity that may be referred to as the 'out of the lab, into the market' approach.

These developments may be read as part of the ongoing 'second quantum revolution', a term coined in the early 2000s (Dowling and Milburn, 2003a), which covers (but is not limited to) technologies emanating from quantum information science. The main subfields of QT are generally characterised as quantum communication/cryptography, quantum computation/simulation, and quantum metrology/sensing. Besides the fact that QT has become a focal point for international investment and competition, why is this field of technology interesting for academic research about innovation? One reason is that while QT and quantum science have been heralded as one of the most important products of human society during the 20th century, with immense practical implications for society and the economy (Jaeger, 2018), exactly what constitutes QT is highly disputed and widely misunderstood. Concepts derived from quantum physics – such as superposition (in which a particle may be in two states, or places, simultaneously), quantum entanglement (whereby classically not-possible correlations can be shared between distant locations), quantum tunnelling (whereby a wave may pass through an ostensibly – according to classical mechanics – impenetrable barrier), or the replacement of absolute truth with probabilistic estimates – make QT not only difficult for people not deeply educated in quantum theory to understand, but also rather difficult to define. The intrinsic difficulty of reaching a widely accepted and precise definition of the concept of QT is amplified when seeking to operationally define QT for the purposes of industry analysis, public investment, or the assignment of intellectual property rights. If a national government wishes to allocate a billion Euro of public funds to QT, what are the actual decision parameters that a public official should use when allocating the funds? If a venture capitalist wishes to invest in QT, what criteria should be used to select appropriate ventures or projects for the firm's portfolio? In an attempt to address these intellectual-cum-practical challenges, we propose in this paper the use of a hybrid human-intelligence/artificial intelligence (AI) approach to 'naturally' defining the domain of QT based on the emergent way that the field has been treated by the world's patent offices in response to inventors in the field seeking exclusive rights for their inventions, i.e., for patents for the technological embodiments of quantum phenomena.

Although several bibliometric studies have been published focusing on the general field and subfields of quantum technologies (Bornmann et al., 2019; Dhawan et al., 2018; Olijnyk, 2018; Pande and Mulay, 2020; Seskir and Aydinoglu, 2021; Tolcheev, 2018), only a few academic studies investigating the patent landscape have appeared (Chang, 2005; Winiarczyk et al., 2013). One possible explanation is that patent data research relying on keyword searches and the use of cooperative patent classification (CPC) codes alone can easily yield 20% 'false positives' (and up to 80% in some cases), with the resulting risk that a significant number of patents belonging to other fields of technology may be falsely attributed to the domain of QT (Travagnin, 2019). To avoid this problem, we have utilised a recently developed supervised machine-learning method to create a new classification tool for quantum technologies to build a cleaner dataset for more accurate analysis.

Patent analysis studies can provide insights into certain aspects of a field such as technological maturity, commercial interest, market formation, and expectations by actors and stakeholders of returns. Furthermore, they can be used as inputs for higher-level analysis, for example on impact assessment for certain policies like industry-academia collaboration incentives. With these factors in mind, and in addition to the general motivation of our study – to employ a hybrid human-intelligence/AI approach to defining QT, focused on the analysis of patents – the research reported here was conducted with two primary aims. First, we aimed to paint a clearer picture of global patenting activities going on in QT, and to draw some preliminary conclusions from our initial analysis, especially on the distribution and the nature of patenting activities in and by the leading countries in the field. Second, we aimed to demonstrate that AI enhanced patent analysis can provide valuable insights about newly emerging fields of technology that are otherwise difficult or impossible to obtain through conventional searches using CPC codes and keywords alone.

2 Methodology

The methodology we employed to carry out our study relied upon three critically important procedures. The first was to develop a formal operational definition of QT based upon a comprehensive analysis of the semantic content of thousands of patents and published patent applications, using the actual substantive characterisation of the technology contained in the patent documents, rather than pre-conceived classifications based on standardised CPC codes or arbitrary or irregular use of keywords. We engaged in an iterative process, in multiple stages, involving the curated application of AI algorithms to gradually train an AI software system to recognise the distinctive features of an invention that made it a QT invention rather than a member of some other field of technology.

This process, a hybrid AI-human process, involved feeding the AI system with pertinent information in the form of patents in the field of QT obtained by well-known actors operating in the QT commercial landscape, and then systematically evaluating examples of QT patents identified by the AI system from a comprehensive global patent dataset covering all patents and patent applications published by patent offices from over 100 countries over the last several decades. In other words, the method required extensive interaction between intelligent humans who 'trained' the AI software to recognise examples of patents in the field of QT, and the AI system itself that presented to the human beings examples it found from published patent documents worldwide that fitted within the parameters of QT patents it had been trained to recognise at each iteration of the training process. The end result of this iterative process was a robust formal 'definition' of QT embedded in to the AI software as a set of characteristic features of QT against which any published patent document – either a patent or a patent application – could be checked. The QT definition, which in effect is a software-embedded algorithm-based classification tool, is labelled for convenience here as a QT 'classifier'. The overall process of training the AI system and producing the QT classifier – which entailed the careful checking of about 11,600 patent documents, and classifying over 5,000 of those documents as describing genuine QT inventions (i.e., as 'positives'), and over 6,000 as failing to do so (i.e., as 'negatives' or false positives) – is described in detail in the Appendix.

Following completion of the first procedure – the creation of a formal operational definition of QT (i.e., the creation of a QT 'classifier') – we then applied the QT classifier to search the whole population of all patent documents worldwide published digitally online, in all fields of technology, to identify which inventions described in the global patent system matched our AI-curated definition of QT. We calculated the cumulative number of active QT patent families worldwide for the most recent 35 years. A patent family was defined as a single set of pending and granted patents worldwide for one invention in which at least one patent or application in the set had been published. Each patent family in our dataset was allocated to a particular year according to the filing date of the first patent application in the family. In this manner we were able to identify over 14,000 patent families worldwide that with very high confidence may be classified as belonging to the domain of QT, conservatively defined.[1] A detailed description of the second procedure is provided in Appendix.

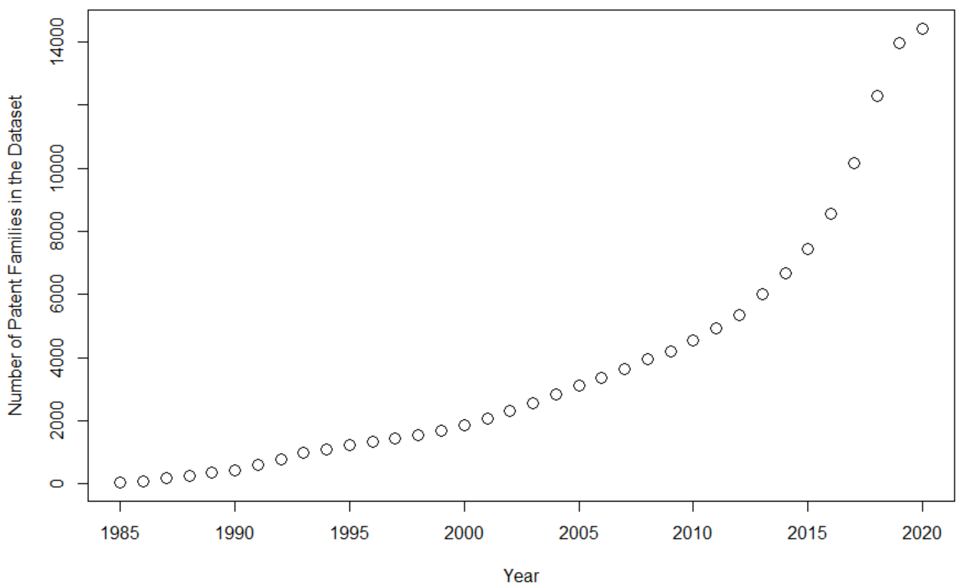
For the third procedure, after having built the final dataset of verified QT patent families during the second procedure, we then applied the AI system to auto-generate technological subcategories from within the dataset itself, by utilising data from the title, abstract, and citations sections of patent documents. In other words, the AI system categorised the broad domain of QT in to sub-fields 'naturally' or endogenously using the internal language of the patent documents themselves – rather than by applying externally imposed classifications such as CPC codes – to characterise the diversity or complexity of the overall QT domain.

All three procedures in our methodology were carried out using a commercially available online patent information service, *Cipher*, developed by Aistemos Limited (2020). *Cipher* includes both the dedicated patent-related AI system itself (accessible as software-as-service over the cloud) and access to a comprehensive curated dataset of all published patent documents (patent applications and granted patents) worldwide for the most recent three-and-a-half decades. Details are provided in Appendix.

3 Profiling the field of QT using an AI-enhanced patent dataset

After completing construction of our robust and verified QT patent dataset, we conducted a series of analyses, commencing with calculating the number of QT inventions embodied in published patent documents each year since 1985.

Figure 1 Cumulative number of QT patent families, by priority year



The basic results, plotted in Figure 1, reveal a generally linear increase in the number of patented QT inventions worldwide during the first two-and-a-half decades of the period included in the graph, followed by almost exponential growth during the most recent decade. The slight levelling in the graph line following 2019 is caused by the fact that most patent applications remain secret for 18 months following their filing date, and not because of a slow-down in the rate of new patent applications. The results confirm that QT is clearly a focus of burgeoning international commercial interest, not just a field of scientific inquiry and laboratory experimentation, and that the level of interest has accelerated during recent years.

Next, as indicated in the methodology section, we used *Cipher's* internal AI-based categorisation facility to create auto-generated technological subcategories. Out of the 12 top QT subcategories, the first five contained the vast majority (93%) of currently active patent families, and hence we focused on these and plotted the number of new patent families in each subcategory with respect to priority years. The results are shown in Figure 2, with the vertical axes representing the number of active patent families in each subcategory and the horizontal axes representing the priority year of each family. All five QT subcategories exhibited significant growth. However, it is notable that inventions in fields related to the subject matter of quantum computing, quantum cryptography, and quantum dot-based applications (mostly sensing) exhibit a dramatically higher number of patent applications than inventions related to single photon sources and photon detection systems.

Figure 2 Annual number of new patent families (by priority year) for the five most dominant QT subcategories

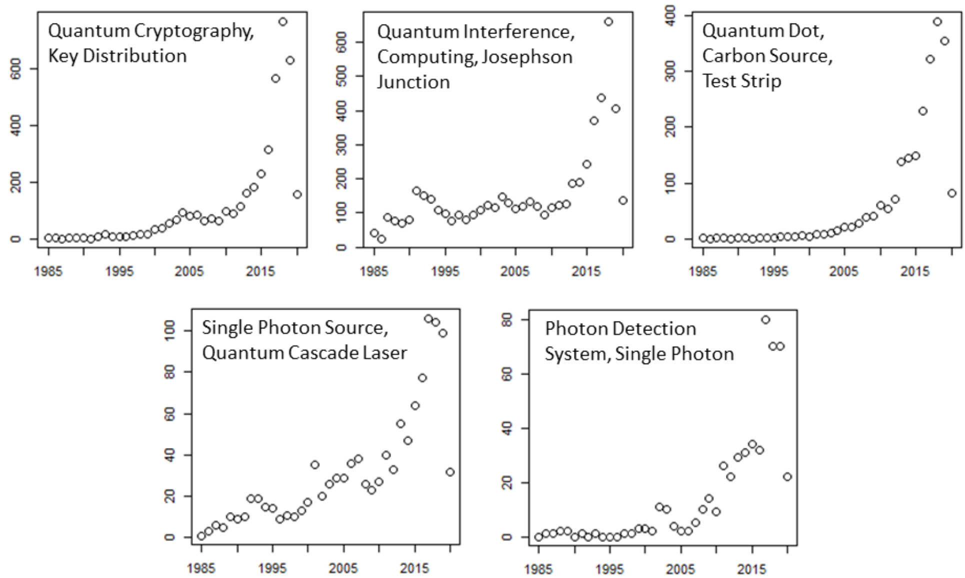
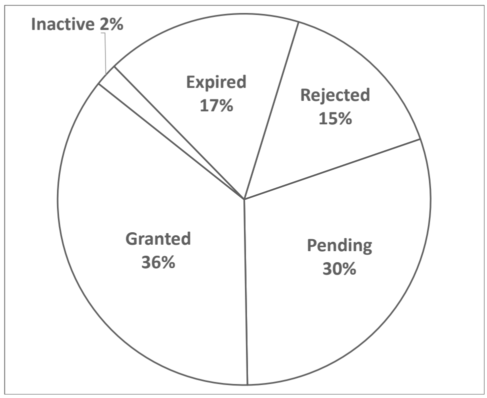


Figure 3 presents a profile of the legal status of the patents in our dataset at the time it was completed (November 2020). In accordance with the protocol of *Cipher*, the basic unit of analysis in our dataset (unless otherwise stated) is the patent family, rather than the individual patent. Thus, if the initial patent application of a patent family (from which the priority date is set) is still pending, the patent family as a whole is registered in the system as pending even if a patent application for the same invention is granted or rejected in other countries. Similarly, if the patent which was the subject of the initial patent application of a patent family is granted, the patent family as a whole is registered as granted, even if a patent application for the same invention is pending or rejected in other countries. Likewise, a patent family in our dataset classified as 'expired' may nevertheless still appear in the results as an 'active patent family' if one or more patents in the family are still active. For example, if an organisation outside the USA applies for a patent in its home country, followed by a patent application for the same invention to the United States Patent and Trademark Office (USPTO), and then subsequently abandons its original application but has its US patent granted, then the patent family to which this invention belongs is classified as 'rejected' even though it contains a granted unexpired US patent.

Our QT dataset includes 2,411 'expired' patent families (meaning that there are 2,411 families which still contain active patents but in which the initial patent in the family has expired). The average time from the priority date to the expiration date of these patent families is 11.8 years, and only 14% of them exhibit less than five years between the priority date and the date of expiration. This indicates that most of the active patent families in our dataset classified by *Cipher* as 'expired' were actually maintained with appropriate payment of fees for quite some time, and hence may be interpreted as having perceived value to their owners. The *Cipher* interface provides an estimated cost-to-date of a patent family in US dollars (US$), and this set of 2,411 patent families together possessed an estimated cost of over US$42 million. The estimated cost-to-date of the whole dataset of all 14,425 patent families is estimated by *Cipher* to be US$214 million, indicating that around 20% of the global spending on QT patent costs has been directed towards patents that are currently 'expired'.

Figure 3 Distribution of patent status in the dataset of 14,425 QT patent families



Our dataset includes a total of 9,535 patent families classified with either a 'granted' or 'pending' status, and which together constitute roughly two thirds of all QT patent families identified by our classifier. As an additional check on the utility and robustness of the classifier, and with the goal of seeing whether the classifier inadvertently included high-profile patents unrelated to QT, we selected the top five patents in our dataset according to their estimated cost-to-date.

Table 1 Estimated cost-to-date of the top five patents in the dataset

Patent owner	*Patent title*	*Priority year*	*Cost to date (US$)*
Dolby Laboratories Inc.	Techniques for using quantum dots to regenerate light in display systems	2009	$425 k
D-Wave Systems Inc.	Analog processor comprising quantum devices	2004	$253 k
Google LLC	Efficient network layer for IPv6 protocol	2013	$246 k
Google LLC	Constructing and programming quantum hardware for robust quantum annealing processes	2014	$238 k
University of California	SQUID Detected NRM and MRI at Ultralow fields	2002	$197 k

We were mildly surprised to discover that the most costly patent in our dataset (owned by Dolby Laboratories) happens to be related to the field of quantum-dot based display systems, since we had not consciously borne that field in mind when training the AI system to build the classifier. Nevertheless, the patent satisfied the 0.7 threshold fit score requirement we had adopted for the study and, upon careful examination of the patent document itself, we concluded that it should indeed be included in the dataset. A quick search revealed that out of 2,714 currently active patent families identified through a keyword search[2] as belonging to the field of quantum-dot based display systems, 170 were included in our dataset. After a manual check of this group of patents – which referred to inventions related to topics such as quantum-dot based LEDs, solar cells, and fabrication methods – we nevertheless decided to retain them in our dataset since for the most part they contained elements that belong within the general domain of QT. We also paid particular attention to the IPv6 protocol patent of Google, which ostensibly appeared unrelated to QT. However, it appeared that through a series of quantum related amendments made to an individual patent (AU2019275673A1) within the 'efficient network layer for IPv6 protocol' family, the family as a whole became associated with QT. For example, as illustrated by the following extract from its abstract, the patent contained several references to known quantum technologies: 'a quantum computational system, comprising an error corrector subsystem in data communication with measurement qubit ...' The other three patents listed in Table 1 are clear examples of quantum technologies. Our review of these five high-cost patents further assured us that our QT classifier is appropriate for reliably identifying patents across a variety of subfields of QT, internationally, regardless of their ostensible provenance or superficial appearance.

4 Insights about innovation in the QT domain

4.1 Comparison of QT patenting in China and the USA

Having confirmed the usefulness of our classifier through a variety of robustness tests, we conducted a comparison of QT patenting in the USA and China, which together hold the majority (55%) of the currently active individual QT patent grants worldwide (followed by Japan at 7%). A more detailed multi-national analysis of the country of origin of national patents is provided below (in Section 4.3), but here we focus on the temporal and quantity differences between the patent applications of the two most prominent countries, namely, the USA and China. Bearing in mind that patents pending are an indicator of likely future 'granted' patents, one of the most interesting differences between these two major players' lies with the relative numbers of their patent applications. China, although becoming an active source of patents in the field only after 2010, currently possess 30% of the active individual patent grants worldwide compared to 25% in the USA. The temporal distribution of granted patents, combined with the high number of patent applications emanating from China (as shown in Figure 4), suggests that the future global balance may shift in the direction of China.

The above results evoke two lines of inquiry. First, how does the number of patents in each country translate into commercial impact? Second, who is actually doing the patenting and where are they doing it? To address the first question, two types of information in our dataset – namely, the type of patents and number of forward citations – may provide useful clues. Both of these measures may be seen as indicators of the quality and commercial potential of patents. China, unlike the USA, permits the granting of utility models, a type of patent with lower patentability standards, lower costs and a shorter time period of protection than normal patents (i.e., than 'invention patents' in China, or 'utility patents' in the USA) (Prud'homme, 2017). One criticism that is often levelled against patenting practice in China, and the rise of Chinese patents, is that they are of lower quality than USA, Japanese or European patents, in part because of the prevalence of utility models and the ease and low cost with which they may be obtained (Dang and Motohashi, 2015; Yang, 2014). One justification for this view is the belief that utility model applications in China are frequently filed not in expectation of commercial return or to protect unique intellectual assets but for other purposes, such as fulfilling KPIs or government mandated project goals. Although a similar argument could in principle also be directed towards US patent applications, the difference between the cost (excluding costs associated with patent searching and patent prosecution) of applying for a Chinese patent (around US$150) (European Commission IPR Hub, 2018) or Chinese utility model (around US$75) (European Commission IPR Hub, 2018) and applying for a US patent (around US$300 at minimum) (USPTO, 2020) is substantial. The difference is even greater when patent attorney costs are taken in to account.

Figure 4 Comparison of granted QT patents and QT patents pending in the USA and China

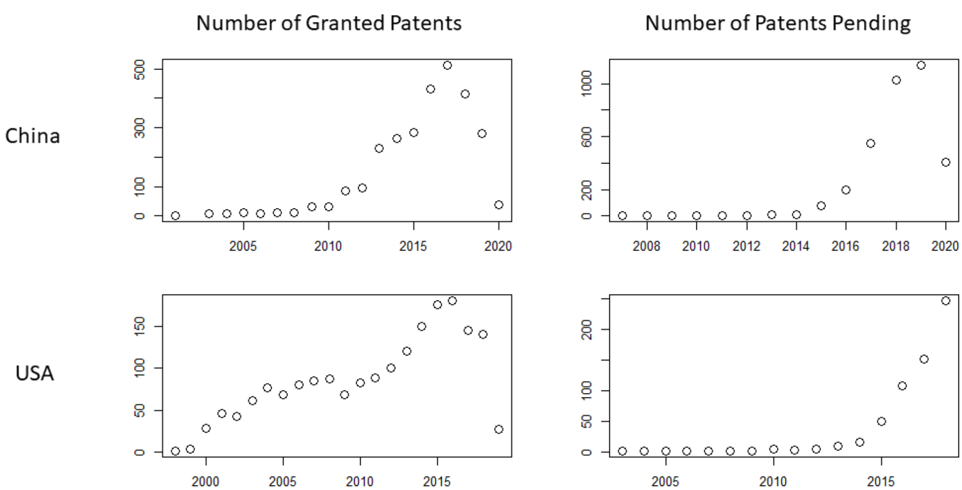


Table 2 Comparison of currently active QT patent families in the USA and China

Patent-related comparison criteria	*USA*	*China*	*China (excluding utility models)*
Number of active patent families	1,870	2,814	2,062
Cost to date [$US] – total	$76 M	$36 M	$31 M
Future cost projection [$US] – total	$45 M	$44 M	$34 M
Average backward citations per patent	19	4	6
Average forward citations per patent	14	3	4
Average priority year	2011	2016	2015

Table 2 enables more subtle comparison of the QT patenting profiles of China and the USA, by presenting two versions of the China data, with one including and one excluding utility models. One interesting revelation that may be gleaned from Table 2 is that utility models account for not much more than one quarter (about 27%) of active QT patent families in China. While this figure is substantial (over 750 QT utility models) it is nevertheless clearly overshadowed by the number of 'real' QT patents. The average number of both forward citations and backward citations per document for utility models is much lower than for normal patents, which indicates that the presence of utility models most likely brings down the overall 'quality' of the national QT patent portfolio in China. Additionally, differences in the average priority date of normal (invention) patents and utility models as shown in Table 2, with the average date for utility models being more recent, suggesting that the proportion of Chinese patents accounted for by utility models may increase in the future, perhaps thereby further lowering the overall quality of the Chinese QT patent stock. This factor may be counterbalanced, however, by the fact that utility models expire earlier than normal (i.e., 'invention') patents in China, and by evidence that the overall quality of patents in China may in fact be increasing over time, due to public policy, corporate awareness and natural learning (Dang and Motohashi, 2015; Prud'homme, 2017; Yang, 2014). Thus, the relative future contribution of China to the QT field may be much more significant than widely presumed.

If one accepts the widespread understanding in the literature that highly cited patents tend to have higher market value than less cited patents (Hall et al., 2005), then it appears from the data in Table 2 that QT patents in the USA tend to be superior to those in China in terms of both quality and market value. This concurs with other published research showing that Chinese-originated patents (in general, across all fields) tend to suffer from a large citation lag in comparison to non-Chinese patents, suggesting a lower value, especially for patents filed domestically (Fisch et al., 2017). However, there is also empirical support for the view that the average value of Chinese patents has been increasing in recent years and that the value gap between Chinese and foreign patents might be narrowing over time (Fisch et al., 2017). Secondly, there is an average of four years of difference between two sets and, as presented in Figure 4, the number of patents pending in China is quadruple the number in the US. At this point, while it appears that the USA outperforms China in terms of the quality of their respective QT patent portfolios, China outperforms the USA in terms of quantity, and there is no guarantee that such simple comparisons may be maintained. The sheer growth in the volume of Chinese QT patent applications, combined with evidence of improvements in patent quality over time, and taking in to account that the vast majority of Chinese QT patent families are based on normal (invention) patents rather than utility models, suggests that the USA lead in quality might not necessarily be maintained in the long term.

4.2 QT patenting behaviour of top organisations

Following the second line of inquiry, we looked at the currently active and granted individual patents per country, for the top 99 QT-patent owning organisations.[3] These 99 organisations together hold 6,257 out of the total of 9,781 active individual patents. We manually assigned each organisation to a country and organisational type. We assigned countries to global organisations with respect to their headquarters. For organisational 'type' we differentiated between four types of organisations:

- quantum start-ups
- established companies expanding into quantum technologies
- academic institutions

- public organisations.

There were several small classification issues we needed to address, such as to which category a limited liability company formed under a public organisation (such as Triad National Security, LLC) should be assigned, or to which country a start-up that moved but retained an office in its original country should be assigned. We classified a company of the first type as a public organisation in our dataset, and we classified a company of the second type as still 'belonging' to its country of origin.

An additional limitation of this part of our analysis is that we counted individual patents instead of patent families, thereby inflating the apparent prominence of companies with inventions patented in multiple countries. For example, Northrop Grumman has 608 individual active and granted patents in our dataset, while having only 130 patents in the USA, meaning that the actual number of QT inventions emanating from Northrop Grumman is closer to 130 than 608. By contrast, IBM has 286 total individual active and granted patents, with 235 of them in the USA. These ratios might provide further insight into the patenting strategies of the companies for later analysis, but for the analysis at hand this is a limitation which needs to be taken into consideration when assessing the findings.

Out of the 99 top organisations as owners of these individual patents, 40 are established companies, 34 are academic institutions, 15 are QT start-ups, and ten are public organisations. First, we looked at the ratio of patents at the home country to total individual patents worldwide, which we labelled the '*Domestic Patent Quotient*' (DPQ).

$$Domestic\ Patent\ Quotient = \frac{\#\ of\ \ patents\ at\ home\ country}{total\ \#\ of\ \ individual\ \ patents}$$

The DPQ may be interpreted as a crude indicator of the international commercialisation intentions of a QT organisation, where a low DPQ indicates a strong orientation towards outward-bound patenting activity, and a high DPQ indicates a stronger focus on domestic patenting.

Table 3 International orientation of QT patenting strategy of organisations

	Companies	*Start-ups*	*Public org.*	*Academic org.*
DPQ	0.35	0.40	0.59	0.68
# of patents	3368	949	469	1,471

Table 3 shows that start-ups and established companies exhibit a higher international strategy orientation than public and academic organisations, with academic organisations being the most domestic in their strategies. The differences between China and the USA in the distribution of patents between the four different types of organisations therefore have implications for their relative overall evolving international strategic positions in QT.

4.3 National comparison of patenting behaviour of top QT patenting organisations

Continuing the analysis, we looked at the national distribution of the 99 top QT patenting organisations. The organisations were distributed across 14 countries, but only four countries were home to more than five organisations in the list, namely, China (31), the USA (27), Japan (11), and Korea (9). Combined, these 78 organisations hold 4,927 active individual QT patents.

Table 4 National variations in the organisational mix of QT patents

	China	*USA*	*Japan*	*Korea*
Percentage of QT patents in each country that are domestic patents	87%	43%	44%	36%
Percentage of QT organisations in each country that are academic organisations	58%	19%	0%	44%
Percentage of QT patents in each country belonging to academic organisations	75%	7%	0%	28%

The ratios presented in Table 4 indicate that organisations based in Korea, Japan and the USA have significantly higher outward-oriented patenting activities than those organisations based in China, which mostly obtain domestic Chinese patents. This finding is related to the fact that nearly 75% of the active QT patents in China are held by academic organisations, whereas the equivalent percentage in the other three countries is significantly lower.

As a final analysis to address our second question – namely, who is doing the patenting, and where are they doing it? We constructed a matrix (see Table 5), wherein the row entries represent the number of active individual patents owned by organisations based in a country, and the column entries represent the number of active patents registered in those countries.

Table 5 International source and destination of QT patents

Home countries of patent owners	*Countries in which patents are issued*								
	China	*USA*	*Japan*	*Korea*	*UK*	*Germany*	*France*	*Canada*	*Total*
China	1,205	54	17	2	10	11	10	1	*1,310*
USA	68	984	115	51	113	103	92	43	*1,569*
Japan	16	186	360	3	79	22	21	2	*689*
Korea	21	108	13	152	11	9	9	1	*324*
UK	6	16	11	2	13	10	10	0	*68*
Germany	2	6	4	1	5	9	5	0	*32*
France	1	21	4	0	18	16	37	0	*97*
Canada	13	189	15	5	15	13	13	17	*280*
Total	*1,332*	*1,564*	*539*	*216*	*264*	*193*	*197*	*64*	*4,369*

Given that the numbers reported in Table 5 cover only active individual patents held by organisations that are in the list of top 99 QT patenting organisations, they by no means provide an exhaustive map, but several key insights are nevertheless evoked. First, the top patenting organisations in China are overwhelmingly oriented towards patenting in China, whereas the major players outside China are much more international in their patenting strategies. Second, as the number of patents from organisations in a country increases, so does the ratio of their domestic patents to foreign patents. Even the top patenting organisations in the USA have more than half of their active patents at home. Canada (or to be more specific, D-wave systems) is the clear outlier here. Third, if we treat the number of patents obtained in a country as an indicator of expected market value in that country, and also accepting that China has an inflated number of national patents due to patents being owned by academic organisations, it is apparent that the USA is the clear expected early market for these technologies.

5 Discussion

Our analysis of the patent landscape in quantum technologies reveals some notable general trends. First, since the early 2010s the growth pattern in the number of new patents in the field has changed from linear to something approaching exponential, and the trend of patenting activities in quantum computing, quantum cryptography, and quantum sensing, in particular, is rising. One particular reason for this increase is arguably the entrance of China into the QT field, as it seems that the proportion of active QT patents worldwide in China rose from near 0% to about 30% in just one decade; and it is worth noting that there are more than 2,000 QT patents pending in China alone (which is near 50% of the total number of patents pending worldwide in the field).

Secondly, we focused on a comparison of China and the USA, the latter being the historical first mover in the field. This revealed that, although the number of active patents in China outweighs the number in the USA, more than a quarter of active QT patents in China are actually utility models, rather than normal ('invention') patents, and that this ratio is on the rise. Additionally, patents in China possess significantly fewer forward and backward citations, and a lower life-cycle cost attached to them. However, the number of patents pending in China is almost quadruple the number in the USA. Thus, we expect that the number of active QT patents that have China as both the country of origin and the destination is going to increase in coming years relative to the USA. Together with these trends, evidence that the quality of invention patents more generally in China has been rising during recent years additionally suggests that China's emerging role as globally prominent innovator and competitor in the world of QT ought to be taken very seriously by both commercial and government actors in the field.

Finally, we moved to investigate the patenting behaviour of top organisations in the QT field and discovered several interesting facts. First, established companies that have expanded into the domain of QT exhibit the highest tendency towards outward-bound international patenting, while academic organisations exhibit the lowest. Second, although the number of organisations in the set does not vary dramatically between companies (40) and academic organisations (34), companies hold more than twice the number of QT patents as academic organisations. Third, the ratios of the types of organisations holding patents vary greatly between countries, as do the ratios of the types of patenting in which they engage. Fourth, the general geographical patenting propensity of the organisations in our dataset is that as the total number of their QT patents increases the proportion that is devoted to domestic patenting also increases (Canada being the outlier). Finally, our analysis revealed that the QT patenting activity of China is overwhelmingly dominated by Chinese academic institutions. This phenomenon makes it difficult to assess the expected commercial value of the huge number of QT patents emanating from China and requires further analysis into the policy decisions and structural factors that led to China becoming the largest national QT patenting actor in the world in just one decade.

6 Conclusions

In this study, we have utilised a new patent classification tool to analyse the global patent landscape of QT. We have provided a detailed description of how this tool was used, and of our tests of its robustness. Furthermore, we demonstrated that using an AI-enhanced patent classification tool may provide substantial insight into both the technical and industrial/organisational characteristics of an emerging domain of technology and that the outcome may be superior to the outcomes of patent analyses that rely solely on CPC codes and keyword searches that, among other things, result in a high number of false positives. As a demonstration of the practical advantages of employing AI-enhanced patent analysis for emerging technological industries, the following insights about the field of QT have been generated: QT is emerging rapidly as a field of technology with substantial industrial and commercial interest (not just academic interest), particularly in the sub-fields of quantum sensing, quantum computing, and quantum cryptography; and, China has a significant and growing presence in the field, as already noted in the literature (Olijnyk, 2018; Sharma, 2018), and is likely to be a major player in the future. However, our research has also indicated that the potential for the emergence of a substantial commercial market for QT – open to international competition – is likely to be seen first of all in the USA, based on the patenting activities of the top QT patenting organisations worldwide.

These findings may guide both private and public actors in the field of QT, and lead to re-alignment of patenting strategies. The findings also evoke several questions for further analysis, including how the various sub-fields of QT and their technological evolution are intertwined; how the core QT technologies are related to complementary technologies; what methods may be available to accurately distinguish and characterise patents from different sub-fields of QT; and, finally, what predictions may plausibly be made regarding future developments in the broad domain of QT based on analysis of previous and current trends in QT patenting activities.

In short, in this paper we have demonstrated how the sophisticated use of intellectual property information, enhanced by the artful deployment of AI techniques, may produce more reliable and useful insights for policy makers and managers about global technological innovation in specific fields than is possible through conventional methods of data collection and analysis.

Acknowledgements

The authors would like to thank Prof. Dr. Jacob Biamonte, Head of the Laboratory for Quantum Information Processing, Skolkovo Institute of Science and Technology, Moscow, Russia, for providing inspiration and advice about recent developments in the fields of quantum information and quantum mathematics.

References

Aistemos Limited (2020) *Cipher*, online patent data service, London, UK [online] https://cipher.ai.

Bornmann, L., Haunschild, R., Scheidsteger, T. and Ettl, C. (2019) 'Quantum technology – a bibliometric analysis of a maturing research field [WWW Document]', *Figshare*, https://doi.org/10.6084/m9.figshare.9731327.v1.

Chang, M. (2005) 'Sun TZU and quantum information – a strategic view for the information technique of tomorrow', in *Proceedings 2005 IEEE International Engineering Management Conference, 2005*, IEEE, St. John's, Newfoundland & Labrador, Canada, pp.730–734, https://doi.org/10.1109/IEMC.2005.1559245.

Dang, J. and Motohashi, K. (2015) 'Patent statistics: a good indicator for innovation in China? Patent subsidy program impacts on patent quality', *China Economic Review*, Vol. 35, pp.137–155, https://doi.org/10.1016/j.chieco.2015.03.012.

Dhawan, S.M., Gupta, B.M. and Bhusan, S. (2018) 'Global publications output in quantum computing research: a scientometric assessment during 2007-16', *Emerging Science Journal*, Vol. 2, https://doi.org/10.28991/esj-2018-01147.

Dowling, J.P. and Milburn, G.J. (2003a) 'Quantum technology: the second quantum revolution. Philosophical Transactions of the Royal Society of London', *Series A: Mathematical, Physical and Engineering Sciences*, Vol. 361, pp.1655–1674, https://doi.org/10.1098/rsta.2003.1227.

Dowling, J.P. and Milburn, G.J. (2003b) 'Quantum technology: the second quantum revolution. Philosophical Transactions of the Royal Society of London', *Series A: Mathematical, Physical and Engineering Sciences*, Vol. 361, pp.1655–1674, https://doi.org/10.1098/rsta.2003.1227.

European Commission IPR Hub (2018) *Patents [FAQs] [WWW Document]*, China IPR SME Helpdesk [online] https://www.china-iprhelpdesk.eu/content/patentsfaqs#:~:text=The basic application fee for, of filing the patent application (accessed 14 November 20).

Fedorov, A.K., Akimov, A.V., Biamonte, J.D., Kavokin, A.V., Khalili, F.Y., Kiktenko, E.O., Kolachevsky, N.N., Kurochkin, Y.V, Lvovsky, A.I., Rubtsov, A.N., Shlyapnikov, G.V, Straupe, S.S., Ustinov, A.V. and Zheltikov, A.M. (2019) 'Quantum technologies in Russia', *Quantum Science and Technology*, Vol. 4, p.40501, https://doi.org/10.1088/2058-9565/ab4472.

Fisch, C., Sandner, P. and Regner, L. (2017) 'The value of Chinese patents: an empirical investigation of citation lags', *China Economic Review*, Vol. 45, pp.22–34, https://doi.org/10.1016/j.chieco.2017.05.011.

Hall, B.H., Jaffe, A. and Trajtenberg, M. (2005) 'Market value and patent citations', *The RAND Journal of Economics*, Vol. 36, No. 1, pp.16–38.

Harris, S., Trippe, A., Challis, D. and Swycher, N. (2020) 'Construction and evaluation of gold standards for patent classification – a case study on quantum computing', *World Patent Information*, Vol. 61, p.101961, https://doi.org/10.1016/j.wpi.2020.101961.

Jaeger, L.A. (2018) *The Second Quantum Revolution: From Entanglement to Quantum Computing and Other Super-Technologies*, Springer, Cham, Switzerland.

Knight, P. and Walmsley, I. (2019) 'UK national quantum technology programme', *Quantum Science and Technology*, Vol. 4, p.40502, https://doi.org/10.1088/2058-9565/ab4346.

Malek, M. (2020) *Personal Communication, Marcus Malek, Head of Cipher Experience*, Aistemos Limited, London, 26 September.

Olijnyk, N.V. (2018) 'Examination of China's performance and thematic evolution in quantum cryptography research using quantitative and computational techniques', *PLOS One*, Vol. 13, p.e0190646, https://doi.org/10.1371/journal.pone.0190646.

Pande, M. and Mulay, P. (2020) 'Bibliometric survey of quantum machine learning', *Science & Technology Libraries*, Vol. 39, pp.369–382, https://doi.org/10.1080/0194262X.2020.1776193.

Prud'homme, D. (2017) 'Utility model patent regime 'strength' and technological development: experiences of China and other East Asian latecomers', *China Economic Review*, Vol. 42, https://doi.org/10.1016/j.chieco.2016.11.007.

Qureca (2020) 'Overview on quantum initiatives worldwide [WWW Document]', *QURECA Quantum Resources & Careers*, 7 September [online] https://www.qureca.com/overview-on-quantum-initiatives-worldwide/ (accessed 14 December 2020).

Raymer, M.G. and Monroe, C. (2019) 'The US national quantum initiative', *Quantum Science and Technology*, Vol. 4, p.20504, https://doi.org/10.1088/2058-9565/ab0441.

Riedel, M., Kovacs, M., Zoller, P., Mlynek, J. and Calarco, T. (2019) 'Europe's quantum flagship initiative', *Quantum Science and Technology*, Vol. 4, p.020501, https://doi.org/10.1088/2058-9565/ab042d.

Roberson, T.M. and White, A.G. (2019) 'Charting the Australian quantum landscape', *Quantum Science and Technology*, Vol. 4, p.20505, https://doi.org/10.1088/2058-9565/ab02b4.

Seskir, Z.C. and Aydinoglu, A.U. (2021) 'The landscape of academic literature in quantum technologies', *International Journal of Quantum Information*, Vol. 19, No. 2, p.2150012, https://doi.org/10.1142/S021974992150012X.

Sharma, M. (2018) 'Decrypting China's quantum leap', *The China Journal*, July, Vol. 80, pp.24–45, https://doi.org/10.1086/697232.

Sussman, B., Corkum, P., Blais, A., Cory, D., and Damascelli, A. (2019) 'Quantum Canada', *Quantum Science and Technology*, Vol. 4, p.20503, https://doi.org/10.1088/2058-9565/ab029d.

Tolcheev, V.O. (2018) 'Scientometric analysis of the current state and prospects of the development of quantum technologies', *Automatic Documentation and Mathematical Linguistics*, Vol. 52, pp.121–133, https://doi.org/10.3103/S000510551803007X.

Travagnin, M. (2019) *Patent Analysis of Selected Quantum Technologies*, European Commission, Joint Research Centre (JRC), Ispra, Italy, https://doi.org/10.2760/938284.

United States Patent and Trademark Office (USPTO) (2020) *USPTO Fee Schedule Effective 2 October 2020* [WWW Document] *USPTO Learning and Resources* [online] https://www.uspto.gov/learning-and-resources/fees-and-payment/uspto-fee-schedule (accessed 14 November 2020).

Winiarczyk, R., Gawron, P., Miszczak, J.A., Pawela, Ł. and Puchała, Z. (2013) 'Analysis of patent activity in the field of quantum information processing', *International Journal of Quantum Information*, Vol. 11, p.1350007, https://doi.org/10.1142/S021974991350007X.

Yamamoto, Y., Sasaki, M. and Takesue, H. (2019) 'Quantum information science and technology in Japan', *Quantum Science and Technology*, Vol. 4, p.020502, https://doi.org/10.1088/2058-9565/ab0077.

Yang, Y. (2014) 'Reforming the utility model system in china: time to limit utility model patents' scope of protection and improve the quality of Chinese utility model patents', *AIPLA Quarterly Journal*, Vol. 42, No. 3, pp.393–424.

Zhang, Q., Xu, F., Li, L., Liu, N-L. and Pan, J-W. (2019) 'Quantum information research in China', *Quantum Science and Technology*, Vol. 4, p.40503, https://doi.org/10.1088/2058-9565/ab4bea.

Notes

1 The actual verified count, as of 12 November 2020, was 14,425 QT patent families worldwide.

2 Search query = ('quantum dot' OR 'quantum dots' OR 'quantum-dot') AND display.

3 We analysed the top 99 QT patent owning organisations rather than the top 100 because the Cipher software groups the patent applications of individuals under the generic label 'Private Owner' and which, when aggregated, technically counts as the 'top' organisation. We therefore excluded the 'private owner' category.

4 Cipher has an in-built scoring system for calibrating the strength of a classifier, on a scale of 0.0 to 1.0, with a higher score indicating greater accuracy in matching the identified patent the to the AI-defined field.

Appendix

Methodology for employing AI to generate an operational definition of QT for constructing a robust global QT patent dataset.

A1 The dataset

A1.1 The patent data: basic features and issues

Our QT dataset was built using *Cipher*, which is a commercially available online patent information service developed by Aistemos Limited (2020) which enables the creation of trained AI classifiers to tag patents against pre-defined technological domains, with the training of the models based on data generated endogenously from the patent records themselves. According to a published study by a team from *Aistemos*, a combination of domain specific normalisation and transformations, and a separately trained patent-specific language model, are used to produce textual and metadata embeddings for the training of models in *Cipher*, and model parameters are obtained through either random or directed hyperparameter searches (Harris et al., 2020). The *Cipher* data base is comprehensive, including data from entire patent families worldwide, and covering both granted patents and published patent applications, which may be searched independently. All training models in *Cipher* utilise information from the title, abstract, citations, CPC codes, and claims of the patent files, but not from the description section of the documents (Malek, 2020). When using Cipher's machine-learning platform it is typically necessary to perform a large number of iterations in the training process to generate representative results (Harris et al., 2020).

To form our QT classification tool ('classifier'), we started with a sample training dataset on 'quantum computing' provided by *Aistemos* and customised it to delete any patents that were clearly unrelated to quantum technologies (e.g., which contained the keyword 'quantum' but obviously contained nothing pertinent to QT), and other patents which consisted mainly of cryptographic schemes 'against' quantum computers (anti-quantum or quantum-resistant schemes). Following this initial filtering exercise, we proceeded to iteratively construct a robust QT classifier by repeatedly refining the training set using the machine-learning tools of the program, verifying the results manually at each step based upon their consonance with the generally accepted scope of the QT field as characterised in the pertinent published literature (Dowling and Milburn, 2003b; Riedel et al., 2019). In short, using the *Cipher* database and software platform, we built an AI-based operational definition (i.e., a 'classifier') of the field of QT using a combination of machine-learning tools and human judgement guided by both the AI and accepted understandings of the technology as promulgated in the pertinent literature.

We faced several issues regarding the construction of a new QT classifier using machine learning techniques:

- As the size of the training set grew larger, re-running the classifier became problematic; leading us to stop the training process at 11,600 patents (5,220 positives and 6,380 negatives). The typical recommended size of a robust the training set is around 750 (250 positives and 500 negatives), but drawing the lines on what can be considered as part of QT and what is not, required further clarification. We did not want the classifier to miss an essential sub-field of the technology just because it did not have enough patents in the training set related to that particular category of technologies.

- Most of the initial patents were selected by hand, following different keyword searches, which meant that the training set contained some internal biases. For example, we tried to include the use cases of quantum dots for sensing and computing purposes while not covering uses such as 'quantum dot based printer ink'.
- One of the main differences between using a trained classifier based on machine learning and keyword searches combined with CPC codes is that for the training set it is necessary to provide not only what 'is' part of quantum technologies but also what 'is not.' In this sense, we believe, depending on the negatives for the training set, a variety of valid classifiers may be constructed, with each focusing on different aspects of QT, and ours is only one example of that wider class.
- Finally, we encountered some grey areas, requiring human judgement, when it came to deciding whether or not certain patent applications should be considered as part of QT. For example, an electronic control system that is designed to be used for registering readouts from a superconducting quantum circuit was accepted as part of QT for the purposes of this study, even though it might plausibly be classified as belonging to a classical system.

Due to the limitations listed above, it is of course possible that a better classifier than ours could be constructed using the same supervised machine learning techniques. Nevertheless, as demonstrated in this paper, the current classifier utilised in this study outperforms analyses conducted using only keyword searches and CPC/IPC codes. Thus, even though our final patent set is non-exhaustive, it is still an accurate and sufficiently large enough sample size to correctly describe the global patent landscape of quantum technologies.

A1.2 Building the dataset and testing its robustness

Initially we tested our classifier by comparing its results with the results of a keyword search provided in literature (Winiarczyk et al., 2013) which is replicated below in Boolean form:

> "quantum computer' OR 'quantum computing' OR 'quantum computation' OR 'quantum compute' OR 'quantum communication' OR 'quantum information' OR 'quantum bit' OR qubit OR qbit OR ('quantum' AND 'random number generator') OR 'quantum cryptography' OR ('quantum key' AND (distribut* OR exchang*)) OR 'quantum Fourier' OR ((quantum OR photo* OR optic*) AND BB84) OR (quantum AND grover) OR (quantum AND ('single photon source' OR 'single-photon source' OR 'single photon generator' OR 'single-photon generator' OR 'single photon detector')) OR (quantum AND spintronic*)."

Out of 4,273 patent families in the dataset generated by the above keyword search, our classifier identified 3,772 patent families above *Cipher's* 0.5 threshold and 3,700 patent families above *Cipher's* 0.7 thresholds for the strength of the patent-classifier-patent fit.[4] We checked the 72 patents that fell between the 0.5–0.7 score by hand, and although half of them definitely belonged to the set of QT (such as single photon sources or quantum cryptographic methods) half of them did not. For example, a patent titled 'A photon source' containing the sentence 'The source may be configured as a single photon source by incorporating quantum dots' was found, which indicates that in its current form and proposition, it is not a single photon source. To root out false positives at the expense of some true positives, we set the threshold as 0.7 for the rest of the study.

We then tested our classifier again by comparing its results with the results obtained by applying an experimental *Cipher* classifier produced by *Aistemos* for the field of 'quantum computing' which contained 4,104 patent families. When we applied our classifier to that reference dataset we obtained 3,721 patent families that satisfied the threshold fit requirement of at least 0.7. Our classifier identified 164 of the 4,104 patent families as falling within the 0.0–0.1 threshold range. Examples of the 'false positives' in the reference dataset included patents with the following titles (plus our explanation of the actual subject matter of the patent):

- 'A kind of quantum wine' (recipe for an actual wine for 'raising sleep quality').
- 'A kind of control system of multi-functional quantum hair tonic comb' (a comb with herbaceous plant essence for nourishing hair and massaging the head).
- 'A kind of quantum generation converting apparatus' (a cup for drinking tea with a leak hole for honey or granulated sugar to be leaked directly into the cup).

To conduct a third round of testing of our QT classifier we built a completely new dataset by casting a wide net to cover as many patents as possible, using a combination of keyword searching and other methods and sources. We commenced with the following broad-scope Boolean search, which contained some redundancies, since it was constructed by combining searches with a selection of keywords drawn from a wide variety of sources, and generated an initial list of around 1.6 million patent families:

> "quantum computer' OR 'quantum computing' OR 'quantum computation' OR 'quantum compute' OR 'quantum communication' OR 'quantum information' OR 'quantum bit' OR qubit OR qbit OR ('quantum' AND 'random number generator') OR 'quantum cryptography' OR ('quantum key' AND (distribut* OR exchang*)) OR 'quantum Fourier' OR ((quantum OR photo* OR optic*) AND BB84) OR (quantum AND grover) OR (quantum AND ('single photon source' OR 'single-photon source' OR 'single photon generator' OR 'single-photon generator' OR 'single photon detector')) OR (quantum AND spintronic*) OR teleportation OR qkd OR qubit* OR 'single photon' OR 'single-photon' OR spintronic* OR entanglement OR qbit* OR entangled OR (gravity AND sens*) OR gravito* OR magnetome* OR nano* OR quantum OR 'quantum simulation' OR 'quantum imaging' OR 'quantum sensing' OR 'quantum sensor' OR 'quantum computation' OR 'quantum computing' OR 'quantum computer' OR 'quantum coding' OR 'quantum programming' OR 'quantum error correction' OR 'quantum error correcting' OR 'quantum circuits' OR 'quantum algorithm' OR 'quantum algorithms' OR 'quantum network' OR 'quantum networks' OR 'quantum channel' OR 'quantum channels' OR 'quantum cryptology' OR 'quantum cryptography' OR 'quantum key' OR 'quantum teleportation' OR 'quantum information' OR 'quantum technology' OR 'quantum technologies' OR semicond* OR supercond*."

In addition to this keyword search, we used *Cipher's* in-built academic patent set covering 151 institutions (mostly universities), which identified around 660,000 patent families. Finally, we added patent families from an additional 278 organisations, by hand, using *Cipher's* suggestion algorithm for organisations with similar portfolios. In total, by combining these three procedures we built a dataset of over 4 million patent families. We then ran our QT classifier against our large manually constructed dataset, using a 0.7 fit threshold, and determined that just 12,149 of the over 4 million patent families in the set actually qualified as belonging to QT according to our classifier. Following that test we then ran our QT classifier against the whole population of all patents, in all technological fields, worldwide (accessible via the platform), using the same 0.7 fit threshold, and determined that 14,425 of all patent families globally could be classified as referring to genuine QT inventions according to our QT classifier (correct as of 12 November 2020). Given that the order of magnitude of the results achieved using our QT classifier against the whole population of patent families worldwide was similar to what we obtained through manual efforts, we decided to use the dataset generated using the AI tools of *Cipher* as the basis for our subsequent analysis of the field.

Finally, the strength of our QT classifier according to the *Cipher* system registered an extraordinarily high score of almost 98% (0.976), which means that we expect and accept an error of between 2% and 3% (around 350 patents) in the form of 'false positives' within the dataset. Table A1 contains the detailed results of the test for the classifier strength.

Table A1 Robustness test for the QT classifier, using cipher algorithm

Test no.	*True positive*	*True negative*	*False positive*	*False negative*	*Precision*	*Recall*
1	1,705	2,059	68	35	0.962	0.98
2	1,698	2,043	84	42	0.953	0.976
3	1,691	2,058	69	48	0.961	0.972
Mean	1,700	2,050	73.7	41.7	0.958	0.976

The precision and recall of the classifier are defined as below, and the classifier strength is given as the mean of the arithmetic average of these two quantities over many repetitions:

$$\text{Precision} = \frac{TP}{TP + FP}, \text{Recall} = \frac{TP}{TP + FN}$$

Following this first robustness test, we tested our QT classifier again, by running it against three independently developed quantum-related datasets from the *Cipher* repository, based on experimental quantum classifiers with a more limited scope. The results, shown in Table A2, in which a high proportion of the patent families in the three comparison datasets are verified as true positives, confirms the robustness of our QT classifier.

Table A2 Robustness tests against three experimental classifiers

Dataset topic – title	*A. No. of purported quantum patent families in comparison dataset*	*B. No. of true positives confirmed by QT classifier*	*Ratio B/A*
Quantum computation	334	325	0.97
Quantum communication	124	114	0.92
Quantum computation	291	284	0.98

Additionally, we tested our QT classifier against patent datasets we constructed in several specific sub-fields of QT using the dominant approach to patent analytics found in the literature, which is to search for patents of interest using CPC codes and specific keywords. Since there are no curated datasets available for the topics of quantum sensing and quantum simulation, and because the choice of appropriate keywords was not obvious, we omitted searches regarding those two subfields. It should be noted, for example, that keyword searches such as 'quantum sensor' reveal a significant number of false positives since phrases like 'light quantum sensor' are widely used for non-quantum type of sensors, and hence utilisation of specific searches on quantum sensing devices is not reliable. We therefore restricted ourselves to constructing new datasets for QT sub-fields in the following four areas: quantum cryptography, quantum computers, quantum communication, and quantum key distribution. The results presented in Table A3 demonstrate that our QT classifier is well-suited to identifying most of the patents in these four sub-fields defined according to CPC codes and commonly used keywords.

Table A3 Robustness tests against four QT sub-field databases

CPC code (and keyword)	*A. No. of purported quantum patent families in comparison dataset*	*B. No. of true positives confirmed by QT classifier*	*Ratio B/A*
H04L9/0852: quantum cryptography	1,792	1,491	0.83
G06N10/00: quantum computers	2,258	1,924	0.85
H04B10/00: (communication) + quantum	88	75	0.85
'Quantum key distribution'	1,251	1,227	0.98

As a final test of the robustness of our QT classifier's ability to identify genuine QT patents, we repeated several keyword searches that were employed in a previous study conducted by the European Commission (Travagnin, 2019), using the Global Patent Index database of the European Patent Office, in which the 'title and abstract (and in case of doubts the entire text with its claims)' were read for each patent application to filter out false positives, thereby identifying 'true positives' with high precision. We then ran our QT classifier against the datasets that we built using those same search terms. The results are presented in Table A4.

The results presented in Table A4 reveal that, on the whole, our QT classifier performs robustly, with a level of precision that is arguably greater than for QT analyses that rely solely on keywords and CPC codes. However, while this may hold true for the dominant fields of QT – namely quantum computation simulation, quantum communication/cryptography, and quantum metrology/sensing – the results appear problematic for cold-atom based applications. This is because these sub-fields of QT were not explicitly included in the classifier when it was built. This should also be noted as a limitation of our dataset. Nevertheless, we may conclude from the results of the robustness tests reported in Tables A1, A2, A3 and A4 that the overall approach of using an AI-based platform for classifying and identifying appropriate technology, and for constructing an instrument for analysing pertinent patents, is a powerful tool for accurately characterising the domain of QT.

Table A4 Robustness tests against results of 2018 European Commission Study

Search query	*Hits (2018)*	*True positives (manual check – previous study)*	*Hits (2020)*	*True positives (our classifier)*
(photon OR photons OR photonic) AND (entangled OR entanglement OR entangling OR entangle)	337	333	475	438
cpc_code: ‘G06N99/002’ OR (((‘qbit’) OR (‘qbits’) OR (‘qubit’) OR (‘qubits’)) OR ((‘quantum computer’) OR (‘quantum computers’)) OR ((‘quantum computation’) OR (‘quantum computations’)) OR ((‘quantum memory’) or (‘quantum memories’)) OR (‘quantum error correction’) OR ((‘quantum simulation’) OR (‘quantum simulations’)))	1,373	1,149	1,630	1,342
cpc_code: ‘H04L9/0852’ OR cpc_code: ‘H04L9/0855’ OR cpc_code: ‘H04L9/0858’ OR (quantum AND key AND distribution) OR (qkd) OR (quantum AND cryptography)	1,338	1,161	3,272	2,782
(spin OR spins) AND (entangled OR entanglement OR entangling OR entangle)	104	24	129	29
(‘cold atom’) OR ((atom OR atoms OR atomic) AND (interferometer OR interferometry))	263	150	391	28