Patenting Patterns in Artificial Intelligence
Identifying National and International Breeding Grounds

Leusin, Matheus Eduardo; Günther, Jutta; Jindra, Björn ; Moehrle, Martin G.

Document Version
Final published version

Published in:
World Patent Information

DOI:
10.1016/j.wpi.2020.101988

Publication date:
2020

World Patent Information 62 (2020) 101988

Contents lists available at ScienceDirect

World Patent Information

journal homepage: http://www.elsevier.com/locate/worpatin

Patenting patterns in Artificial Intelligence: Identifying national and international breeding grounds

Matheus Eduardo Leusin [a,*], Jutta Günther [a], Björn Jindra [a,b], Martin G. Moehrle [c]

[a] Faculty of Business Studies and Economics, Chair of Innovation and Structural Change, University of Bremen, Germany
[b] Department of International Economics, Government and Business, Copenhagen Business School, Denmark
[c] Institute of Project Management and Innovation, University of Bremen, Germany

ARTICLE INFO

Keywords:
Artificial intelligence
Digitalisation
Patent analysis
International comparison
Specialisation

ABSTRACT

This paper identifies countries at the forefront of Artificial Intelligence (AI) development and proposes two novel patent-based indicators to differentiate structural differences in the patterns of intellectual property (IP) protection observed for AI across countries. In particular, we consider (i) the extent to which countries specialise in AI and are relevant markets for corresponding IP protection ('National Breeding Ground'); and (ii) the extent to which countries attract AI from abroad for IP protection and extend the protection of their AI-related IP to foreign markets ('International Breeding Ground'). Our investigation confirms prior findings regarding substantial changes in the technological leadership in AI, besides drastic changes in the relevance of AI techniques over time. Particularly, we find that National and International Breeding Grounds overlap only partially. China and the US can be characterised as dominant National Breeding Grounds. Australia and selected European countries, but primarily the US, are major International Breeding Grounds. We conclude that China promotes AI development with a major focus on IP protection in its domestic market, whereas the US sustains its AI progress in the international context as well. This might indicate a considerable bifurcation in the structural patterns of IP protection in global AI development.

1. Introduction

The transformation from an analogue into a digitalised world economy has been underway for some decades, manifested primarily by the diffusion of information and communication technologies into the realm of business and society. In the current wave of digitalisation, the development and use of Artificial Intelligence (AI) represents a qualitatively new development. Despite its more than 60 years of existence, AI's potential and market relevance has been recognised worldwide only in the last decade, mainly due to the development of high-performance parallel computing chips and large datasets that have extended this technology's applicability [1,2]. Nowadays, AI can be embedded in any technology (software, algorithm, a set of processes, a robot, etc.) that is able to function appropriately when endowed with the foresight of its environment [3].

Given its scope and potential impact, AI is currently considered a strategic technology for many countries. Accordingly, a 'global AI race' was deployed in pursuit of its development, with countries increasingly investing in national AI strategies intended to gain advantages over global markets and industries [2]. However, it is known from Innovations Systems (IS) theory that, even if countries focus their efforts on the same direction, distinct national characteristics will affect technological development [4]. Beyond national boundaries, it is also known that a technology, or the knowledge it embodies, is rarely embedded in just the institutional infrastructure of a single nation or region, since the relevant knowledge base for most technologies originates from distinct geographical areas [5,6]. At the same time, the intrinsic characteristics of each type of technology affect both the development of their particular body of knowledge and the diffusion of its applications.

To understand the particularities of the 'global AI race', we want first to identify which countries have been leading this development. Moreover, we agree with [5] in the sense that, once innovation systems co-evolve with the process of technological change, it is important not only to identify a final and static picture of the technological development but the dynamics of its evolution over time. Thus we analyse, over time: i) which countries have been specialising in and increasingly adopting AI technologies; ii) which countries seek to enhance the attractiveness of their markets to foreign inventors and companies by

* Corresponding author.
E-mail address: leusin@uni-bremen.de (M.E. Leusin).

https://doi.org/10.1016/j.wpi.2020.101988
Received 18 July 2019; Received in revised form 1 August 2020; Accepted 2 August 2020
Available online 30 August 2020

establishing strong local protection of AI-related Intellectual Property (IP) and/or seek to protect their own AI-related IP under the applicable laws in foreign markets; and iii) which corresponding trends we can observe for specific AI techniques across countries.

For this purpose, we use worldwide patent data and adopt a very broad concept of what an AI patent is: besides searching for all patents with some direct reference to AI, we also search for AI techniques, which are advanced statistical and mathematical models used to implement AI functions such as vision, language, and decision-making. We use established indices to analyse the technological specialisation of countries and, as a novel contribution, we introduce two indexes: the National Breeding Ground (NBG) and the International Breeding Ground (IBG). The first orders countries on the basis of market favourability, that is, it assesses the extent to which the market is perceived as being exploitable by inventors and companies, no matter where they come from. The second orders countries on their ability to attract applications for domestic AI patents from foreign inventors/companies, as well as the degree to which AI patents registered in the country provide legal protection abroad.

We find that the number of AI-specialised countries has grown in the last decades, with a more pronounced growth seen mainly in the 90's. Despite this, the exploitation of AI-related techniques became much more intense after the turn of the century, pushed mainly by the use of biological models (neural networks, supervised and unsupervised machine learning, deep learning, etc.). In terms of international leadership, we can confirm a decline in terms of relevance of Japan and some European countries, conversely to a greater increase in the relevance of the US and the emergence of China. These two leading countries manage AI exploration very distinctly, as described and explained in the next sections.

2. Literature review: Breeding ground indicators and related literature

Among the many indicators used in the literature for comparing country performances in e.g., FDI, number of patents produced, or amount of R&D spent at the national level, the Revealed Comparative Advantage (RCA) index stands out [7]. Introduced in [8], the RCA indicator was proposed as a measure of the *relative* specialisation of a country in the production of a specific product. The simple idea behind the indicator is that when a country has a production level higher than the global average, this country holds a comparative advantage in producing it. Eq. (1) presents the formula for calculating the RCA of a given country '*b*' for a specific product '*a*'.

$$\mathrm{RCA}_{\mathit{Product}\ \ a,\mathrm{Country}\ \ b} = \frac{\frac{\text{Exports of product } a \text{ in country } b}{\text{Total of exports of country } b}}{\frac{\text{Total exports of product } a \text{ in the world}}{\text{Total exports in the world}}} \tag{1}$$

The RCA has been adapted to analyse, among others [9] the comparative advantage of a given country for a specific technology/technological field, through the so-called Revealed Technological Advantage (RTA) index. Like the RCA, the RTA also has extensive application in the literature, with a recent example in Weresa [10], in which the author analyses the comparative advantages of the European Union in digital technologies using patents as inputs for this index.

However, despite its popularity, the RCA (and its related adaptions, as the RTA) has some limitations, including the challenge of interpreting values above 1, highly asymmetric distribution, and the consequent impossibility of directly interpreting the resulting RCA values [11]. The literature [9] points to the prominence given by the RTA to less patent-active countries, and to the problem of statistical bias when the overall number of patents in a specific technology/technological field is too small. In this way, countries with a small national output (e.g., technological output as patents) tend to be highly favoured in cross-national comparisons. Namely, an above-one RTA value signals a country's specialisation, even if the country's absolute output is low. However absolute outputs are highly relevant for market comparisons, especially when discussing the adoption of a new technology: to attract relevant inventors and companies involved with the technology, a country must offer not only some kind of specialisation, but also a relevant market.

To assist technology holders with evaluating competing marketplaces and to enable countries to better understand where they stand, we propose 'Breeding Ground' indexes. The basic premise behind them, supported by qualitative discussions in innovation system theory, is that particular characteristics of each innovation system affect innovators with regard to the markets they choose to exploit their technologies. Accordingly, we aim to identify countries that tailor their National Innovation System towards the adoption of a given technology by providing specific local conditions. In our case, the first favourable conditions considered are the availability of local specialisation advantages together with an attractive domestic market to exploit the technology, which we measure jointly with the proposed NBG index. We expect that countries with a higher NBG number can explore this characteristic to develop the so-called national champions, as highlighted for example in Kroll [12] in an analysis of China, which points that, powered by a strong domestic market, such companies might become competitive in foreign markets. In addition, we propose the IBG index, which we use to identify countries that are attractive to inventors/companies from abroad while also having their own national AI-patents recurrently extended to markets abroad (i.e., by the extension of their patents' legal protection to foreign countries). We argue that this index reflects not only the relevance (size) of markets for AI exploitation originating from abroad, but also efficient institutions reinforcing IP protection and cooperative behaviour [13]. Furthermore, the indicator reflects the national production of AI technology considered relevant for protection on other foreign markets. A high country score on the NBG index but not on the IBG index might signal that strong technological development and exploitation, as well as corresponding patterns of IP protection, take place mainly in a closed domestic arena rather than in an international context.

3. Data collection

In 2019, WIPO [14] argued that despite the availability of information in patent documents, it can be difficult to identify exactly which patent families relate to AI because of the lack of a standardised definition; even non-standard definitions of AI change over time. The literature proposes a variety of strategies for identifying AI-related documents (e.g., patents or publications), including the use of predefined classes based on patent classification schemes [15,16], the use of specific keywords [17–20], or even both [14,21,22]. Both strategies have pros and cons: the choice of keywords and IPC codes is inherently subjective – the first choice depending on which keywords are considered relevant and that of patent officers depending on the interpretation from IP specialists on the content provided in the application as they seek to classify them (see for example [23] for a recent related discussion for the use of IPC codes). Thus, our intention here is not to extend this discussion in the sense of defining AI precisely (even the syntagm 'Artificial Intelligence' has only recently reappeared in industry, after some hypes and disillusionments, as noted in [24]). Rather, with the aim of developing an overview of AI patenting activities worldwide in mind, we seek to create a dataset that is strongly related to the core of AI development through the years. For this, we collected data on all patents with some direct reference to Artificial Intelligence or that are based on typical AI techniques.

To identify relevant AI techniques, we used the framework proposed in WIPO [14], which includes the analysis of AI experts on both the applied and the research domains on this topic. This framework, based on a computing classification scheme developed over the past 50 years by the Association for Computing Machinery, differentiates three main categories related to AI: i) AI techniques as advanced forms of statistical

Table 1
AI techniques considered and definitions.

AI Technique	Definition	Additional wikipedia synonyms
Bio-inspired approaches	A family of AI approaches inspired by biological systems, rather than a precise technique.	Bio-inspired computing, biologically inspired computing
Classification and regression trees	Predictive models that use tree-like representations of facts and their possible consequences.	Decision tree learning
Deep learning	A machine learning approach that tries to understand the world in terms of a hierarchy of concepts. Most deep learning models are implemented by increasing the number of layers in a neural network.	Deep structured learning, hierarchical learning
Description logistics	A form of programming used in Logic programming.	Keyword not found
Expert systems	A computer system that solves complex problems within a specialised domain, based on an expertise expressed manually by human experts in the form of a set of rules.	No additional synonym
Fuzzy logic	A decision-making approach that is not based on the usual 'true or false' assessment, but rather on 'degrees of truth' (where the 'truth' value ranges between completely true and completely false).	No additional synonym
Instance-based learning	A family of machine learning algorithms that compare a new problem with cases seen in training and can adapt the model to previously unseen data.	Memory-based learning
Latent representa-tion	The mathematical representation of variables that are inferred rather than directly observed. Latent representation is applied in natural language processing, for example, where it is usually inferred from the statistical distribution of words.	No additional synonym
Logic programming	Uses facts and rules to make decisions, without specifying additional intermediary steps, in order to achieve a particular goal.	No additional synonym
Logical and relational learning	It is a form of learning related to Machine Learning.	No additional synonym
Machine learning	An AI process that uses algorithms and statistical models to allow computers to make decisions without having to explicitly program it to perform the task.	No additional synonym
Multi-task learning	A machine learning approach where a single model is used to solve multiple learning tasks at the same time, exploiting commonalities and differences between the various tasks.	Multitask Learning
Neural networks	A learning process inspired by the neural structures of the brain, being the network generally organised in successive layers of functions, with each layer using the output of the previous one as an input.	No additional synonym
Ontology engineering	A set of tasks related to the methodologies for building ontologies, namely the way concepts and their relationship in a particular domain are formally represented.	No additional synonym
Probabilistic graphical models	A framework for representing complex domains using distribution of probabilities; the models use a graph-based representation for defining the statistical dependence or independence relationships between data.	Graphical model, structured probabilistic model
Probabilistic reasoning	An approach that combines deductive logic and probability theory to model logical relations under uncertainty in data.	Probability logic, probabilistic logic
Reinforced learning	An area of machine learning that uses a system of reward and punishment as it learns how to attain a complex objective.	Reinforce-ment learning
Rule learning	Machine learning methods which identify and generalise automatically a set of rules (which are usually simple conditional tests) to be used for prediction or classification of new, unseen data.	Rule induction
Supervised learning	The expected grouping of the information in certain categories (output) is provided to the computer through examples of data (input) that have been manually categorised correctly and comprise a training dataset. Based on these examples of input-output, the AI system can organise new, unseen data into predefined categories.	No additional synonym
Unsupervised learning	A type of machine learning algorithm that finds and analyses hidden patterns or commonalities in data that has not been labelled or classified.	No additional synonym
Support vector machines	A supervised learning algorithm that analyses labelled/grouped data, identifies the data points that are most challenging to group and, based on that, identifies how to separate the different groups and classify unseen data points.	Support vector networks

Source: WIPO [14] (pp. 148–150).

and mathematical models; ii) AI functional applications; and iii) AI application fields (i.e., areas or disciplines in which AI techniques or functions may find application). In our investigation, we analysed AI techniques as the technological core that enables diffusion into related functional applications and broader application fields. Thus, we adopted the 21 keywords related to AI techniques proposed in the WIPO report (see [14], p. 24), complementing these with their synonyms (to avoid losing relevant patents due to different wording) collected from Wikipedia. In addition to being the largest knowledge repository of the Web, Wikipedia offers multi-faceted and cross-linked classifications and concepts [25]. [26] further emphasises Wikipedia as the best available option for gaining a comprehensive understanding of the many technical terms used in patents. The complete list of keywords and synonyms considered is presented in Table 1, together with the techniques' definitions given in [14].

Building on our previous definition, while some AI patents are related to the use of AI techniques, others reference AI directly. Hence, we also included the search term 'Artificial Intelligence' and its Wikipedia synonym 'Machine intelligence'.

The patent search was conducted in the autumn 2017 version of PATSTAT (PATSTATb). We identified 40,481 patent applications[1] (each of which has a unique application ID,[2] 'appln_id') whose title or abstract contains the above-outlined keywords. The full query used for retrieving these Application IDs is presented in Appendix A. Once the application IDs were identified, the remaining relevant data was retrieved from PATSTAT using these applications IDs as input.

As a test of robustness, we identified the number of patents related to each combination of the terms adopted within the collected sample. These results, as well as the exact keywords used for the collection and for this robustness analysis, are presented in Table 2.

The inverse check resulted in 42,736 patents, which implies that at

[1] An application is a request for patent protection of an invention. Applications are registered on PATSTAT whether or not they have been granted.

[2] The appln_id is a numerical technical identifier used in all PATSTAT databases to uniquely identify a patent application, allowing the identification of the same application across all editions of all PATSTAT databases [27] EPO, Data Catalog – PATSTAT EP Register – 2018 Autumn Edition, 2018.

Table 2
Keywords used for the patent search.

AI technique keyword	Additional wikipedia synonym keyword	No. of patents
%neural network%	No additional synonym	19,784
%machine learn%	No additional synonym	5228
%artificial intelligence%	%machine intelligen%	4197
%expert system%	No additional synonym	3838
%support vector machin%	%support vector network%	3442
%fuzzy logic%	No additional synonym	2883
%graphical model%	%structured probabilistic model%	806
%pervised learn%	No additional synonym	667
%deep learn%	%deep structured learn% OR % hierarchical learn%	663
%classification tree% OR % regression tree%	%decision tree learn%	415
%reinforced learn%	%reinforcement learn%	375
%logic programming%	No additional synonym	152
%rule learn%	%rule induction%	111
%probabilistic reason%	%probability logic% OR % probabilistic logic%	60
%task learn%	No additional synonym	56
%logical learn% OR % relational learn%	No additional synonym	30
%latent represent%	No additional synonym	10
%bio-inspired approach%	%bio-inspired comput% OR % biologically inspired comput%	7
%instance-based learn%	%memory-based learn%	7
%ontology engineer%	No additional synonym	5
%description logistic%	Keyword not found	0

Note: The double characters '%' at the beginning and end of each keyword are used to include variations before or after this character so that any patent that coincides with the term between these two characters is collected. The search in PATSTAT is not case-sensitive.

least 2255 patents (or approximately 5.6% of the sample) have a combination of at least two of the selected terms. One might expect that keywords like "graphical model" would be too generic, but it is seen that this broad keyword generated only 860 (1.9%) of the patents. At the same time, 92.1% of the patents found concentrate around 6 of the 21 terms used (namely Neural networks, Machine learning, Artificial Intelligence or Machine intelligence, Support vector machines or Support vector networks, Expert systems, and Fuzzy logic).

For the final dataset, we excluded the patents of utility models. This is because, besides having a shorter protection terms and grant lags, utility models have been used by IP professionals as auxiliary tools in specific national contexts to overcome shortcomings of the patent system, as discussed in [28]. Furthermore, we also consider the distinction between priority and non-priority filings. In short, a priority filing (or priority patent) is the first patent application filed to protect an invention. It represents the total number of patent families, regardless of their spatial protection scope. After a priority filing, if the same patent is registered in other patent offices, the following registrations are called non-priorities, constituting a patent family linked through the priority filing. In PATSTAT, an application ID (appln_id) enables the retrieval of information about the first filling ID (earliest_filing_id) associated with this patent. If the application ID and the first filling ID are equal, this application ID is considered a priority filing; if not, this particular application ID is considered a non-priority filing.

We also took into account the distinction between registrations made under the Patent Cooperation Treaty (PCT) and standard registrations filed only with national patent offices. The PCT, an international patent law treaty, provides a unified procedure for filing patent applications to protect an invention in each of its 152 signatory states [29,30]. The PCT registry doesn't grant nor examine patent applications. Instead, it allows the applicants of a patent to delay the expensive step of filing other foreign patent applications. Thus, the benefit of such PCT route is to allow applicants to seek patent protection simultaneously in a large number of countries. The possible following registers of the patent will have the filing date of the first application, and they cannot be invalidated through any acts that occurred during the interval allowed by the PCT [31]. Furthermore, the PCT registry also extends the period to which a subsequent filing can be registered based on the priority filing from the regular 12 months–31 months; in this way, the applicant has more time to assess the potential of its invention and proceed or not with the patent application [32]. [32,33] point out that although some bias might exist, empirical evidence suggests that the PCT route is associated with higher-value patents (or at least, inventions with high market potential, as also argued in [34]). Accordingly, we assume that PCT registrations are chosen by assignees for the patents they consider more valuable on international markets, thus contributing more than other registrations to the international commercialisation of AI.

4. Comparison of strategies for identifying AI patents

As already mentioned, there is a wide variety of possibilities for searching patents on a given topic, which includes the use of keywords and the use of patent classification schemes. Given our use of a keyword-based strategy, we compared our results with those from two other AI-related papers [15,16], that are instead based on the use of codes from the International Patent Classification (IPC).[3] The mentioned authors use a search based on the IPC codes presented in Appendix B.

In short, every code used by [16] pertains to the subclass 'Computer systems based on specific computational models' (G06 N), while [15] also includes codes related to the subclasses 'Optical Computing Devices' (G06E) 'Analogue Computers' (G06G), 'Hybrid Computing Arrangements' (G06J), 'Electric Digital Data Processing' (G06F), and one related to the subgroup of electric adaptive control systems (G05B 13/02). Excluding utility models for the sake of comparability, the search from [15,16] results in 23,599 and 146,049 priority filings, respectively. To compare the efficiency of these search strategies with ours, we manually selected the first[4] 100 patents from our dataset that i) had a title and abstract; and ii) were written in English. We then selected the first 100 records from the two competing datasets that, in addition to our criteria, iii) were not in our dataset. This yielded a total of 300 patents, which we classified individually based on title and abstract.

Although our intention here is not to extend the discussion in the sense of defining what AI is, we do need to define what we consider an AI patent to make this classification possible. Thus, considering our broad view on AI, we classified as AI-related all patents that met at least one of the following criteria: i) can be used to generate some kind of prediction or classification useful to make some decision, or to interpret or summarise some type of knowledge; ii) enables the automation or optimisation of some task or parameter used in the patent to perform or improve some kind of selection; iii) enables the generation of useful and analysable data, or the autonomous correction of existing data; iv) is related to some kind of training, learning or dynamic adaptation based on data; v) enables the recognition or evaluation of objects or patterns of interest.

The results of this comparison are available in our public GitHub repository,[5] which also includes all data used in this paper and the associated R codes for reproducing it. In total, 13 patents in the sample (6 from our Query, 6 from Query 2 and 1 from Query 3) did not have enough information for being classified with certainty, so we categorised them as 'Unclear'. Fig. 1 summarises this comparison, highlighting the number of results that overlap in Queries 2 and 3 in relation to our query, and the accuracy of each query in relation to the total of

[3] Available in: https://www.wipo.int/classifications/ipc/ipcpub.

[4] First here means that, for reproductive purposes, each sample of patents was selected in ascending order in relation to appln_id.

[5] https://github.com/matheusleusin/Patenting-Patterns-in-Artificial-Intelligence.

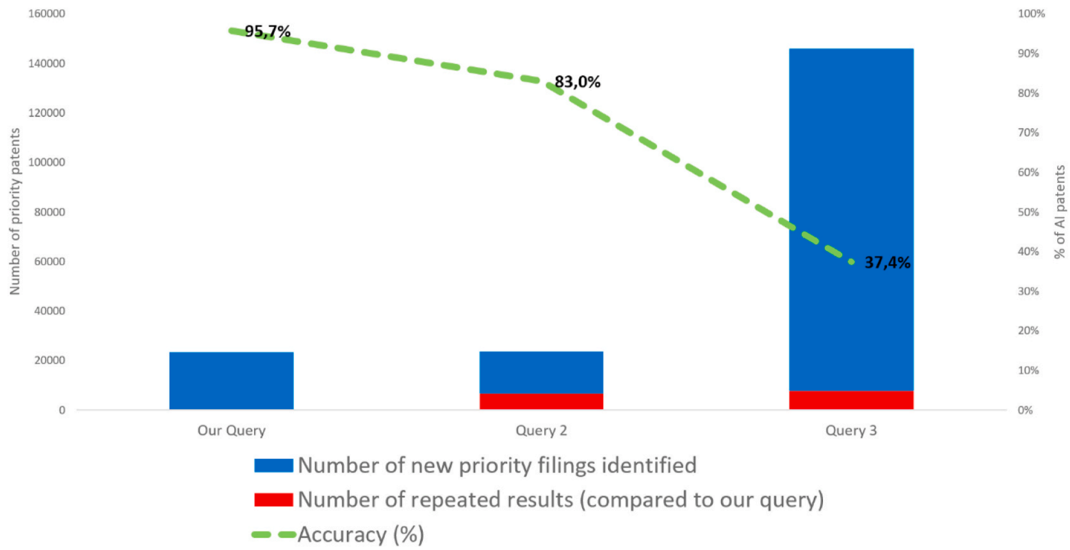

Fig. 1. Comparison of results between the selected queries. Source: Authors' elaboration.

287 records analysed in what concerns being an AI patent or not.

It can be seen from Fig. 1 that there is a significant overlap of results between the two queries considered in relation to ours: 28.2% and 33.1% of our results are also found using Queries 2 and 3, respectively. However, considering the mentioned definition of AI patents, those two queries are significantly less capable of detecting such patents. Query 3, in particular, expands greatly the number of results by including IPC codes less related to AI, which drastically reduces the quality of results generated by analysis of this dataset (AI patents comprised only 37.4%). Our query also performs significantly better than Query 2 (95.7% accuracy compared to 83.0%), thus generating a similarly sized data set of higher quality. One possible explanation for these outcomes, discussed in [23], is that IP offices tend to classify patents by their own standards, potentially losing important features of the invention in the translation into standardised IPC language. The authors even highlight that computer technology is one of the technical fields that demands especial attention by researchers who consider applications across IP offices, because patent applications are often relevant to a variety of classifications and quite-similar applications are coded differently in distinct offices. Thus, the use of a keyword-based search is shown to be a good choice for circumventing the challenges faced by other approaches to gaining a global perspective on far-reaching and still-emerging technologies like AI. In particular, the use of keywords related to AI techniques increases the probability that the patents found are associated with purposes typical of AI, which further improves the quality of the dataset.

5. Method and results

The patent dataset was separated into three periods of twelve years, thus excluding patent applications prior to 1979 (17 patents) and – since the version used of PATSTAT does not show the complete patents of the years 2016 and 2017 – applications after 2015 (4662 patents), resulting in 34,679 priority and non-priority filings to be used as input for calculations related to the specialisation values, and to the attributes of National and International AI Breeding Grounds, discussed in the following sections. The basic idea of our method is using established indices for international comparisons, but developing novel combinations of them, as done in technology management to measure a technology's performance [35].

Before calculating these indexes, a choice had to be made regarding the identification of the countries of origin of the patents. Classically, there are two types of information that can be used for this purpose: the country of origin of the patent inventor, or the country of origin of the patent applicant [36]. Other possible sources, used in [32] when more direct information was not available, is the country of the patent office where the priority filing was registered. The use of this type of information is especially suitable for our analysis: patent offices enable patent protection in the country where they are located (with some exceptions such as the European Patent Office (EPO), the International Bureau of the WIPO, and the Eurasian Patent Organization (EAPO), for example, which enable protection across more than one country), and are therefore a relevant proxy for identifying the markets that inventors and companies prioritise when applying for patents. Thus, we chose to define the origin of each patent as the country whose patent office registered the priority filing related to the patent. A disadvantage of this choice is that the country of origin cannot be properly identified when the priority filing is registered in patent offices that cover more than one country, such as the EPO, WIPO or EAPO. Consequently, we may underestimate the performance of countries covered by wide-ranging patent offices.

Once the country of origin of each patent application is defined, the indicators used in this paper could be easily calculated. To do so, we relied on the variables presented in Appendix C. Our subsequent analysis is divided into 3 steps (see Fig. 2).

First, we referred to the well-known RTA index to analyse the specialisation of countries in the course of time. Second, we used the RTA index calculated on the previous step as one of the inputs of our indicator for National Breeding Grounds, which combines a country's relative extent of specialisation (RTA value) with the absolute number of AI patent applications (unweighted and weighted) filed in that country. Finally, we calculated the International Breeding Ground index, which takes into consideration the country's number of patents and their IP extension abroad, as well as the number of foreign patents registered in this given country (also considering unweighted and weighted patents).

5.1. AI specialists by revealed technology advantage

We use the RTA index to measure the specialisation advantages of countries. An index result of 0 indicates that the country has no patent in the sector considered, 1 when the company's share in the sector equals its share in all fields, and above 1 when the country has a positive specialisation in the sector. We calculate the index for three 12-year periods (see Eq. (2)).

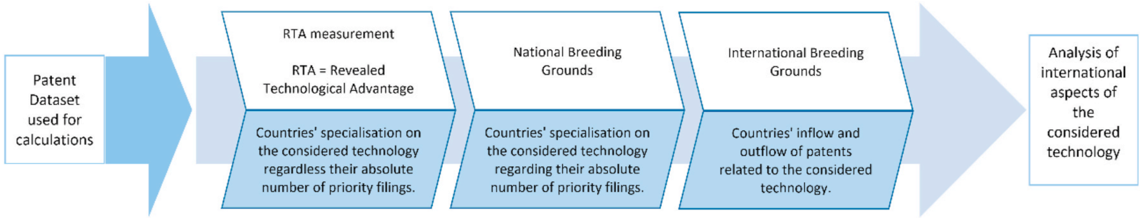

Fig. 2. Description of the steps followed. Source: Authors' elaboration.

$$\text{RTA\_Country}_p = \frac{1}{n}\sum_{t=1}^{n}\frac{\dfrac{\text{Priority Patents AI\_Country}_{t,p}}{\text{Total Patents\_Country}_{t,p}}}{\dfrac{\text{Global Number of AIPatents}_{t,p}}{\text{Global Number ofPatents}_{t,p}}} \tag{2}$$

The variable 'Priority Patents AI_Country$_{t,p}$' is defined as the number of AI priority filings whose application authority is that of the country considered during year t, which is in period p. Similarly, the variable 'Total Patents_Country$_{t,p}$' represents the total number of priority filings registered by the application authority belonging to the country considered in year t and in period p. The 'Global Number of AI Patents$_{t,p}$' and the 'Global Number of Patents$_{t,p}$' represent the total number of AI priority filings registered and the total number of all priority filings registered, respectively, both at the global level, in year t and period p. As previously stated, t varies between 1 and 12, and p between 1 and 3.

The results for the 20 countries with the highest RTA values for each of the considered periods are presented in Fig. 3. The three vertical lines indicate the average RTA value of these 20 countries for the periods considered.

Most remarkably, the number of patent offices with specialisation in AI technologies increases over time, going from two to ten and twelve patent offices with an RTA greater than 1 in periods one, two and three, respectively (Fig. 3). The average RTA values increase from one period to the next, although this increase is greater between periods one and two than between periods two and three.

Furthermore, two main clusters are identified concerning the evolution and RTA values: a European cluster and an Asian cluster. In the European cluster, the Eastern and Southeastern European patent offices (Serbia, Romania, Lithuania and Belarus and Greece) stand out, all sharing the characteristic of having achieved specialisation values only in the third period. The other two patent offices of the European cluster, United Kingdom and France, still present no specialisation. Patent offices of the European cluster still share the characteristic of having increased their RTA values over time, with the exception of France, which is the only patent office in this cluster whose RTA declined in the most recent period.

The Asian cluster includes three out the four Asian tigers (Singapore, South Korea and Taiwan). Together with Malaysia, these countries share the characteristic of having reached the peak of their specialisation in the 1990s, at the pinnacle of their industrial spurt, decreasing in the following period. Japan and China also belong to this cluster, with very different patterns: Japan was an early specialist in AI patents, and was a leader in the first period (together with the EPO and the US), but loses its specialisation in the following two periods, while China achieves specialisation status in the third period.

Finally, among the remaining patent offices, Canada and the US stand out. Both present very similar patterns of evolution, with their RTAs increasing from period one to period two and, surprisingly, decreasing in period 3. Israel and New Zealand also show a decrease in their RTA values in the third period. Patent offices that facilitate the registration of a patent for more than one country, like the EPO and

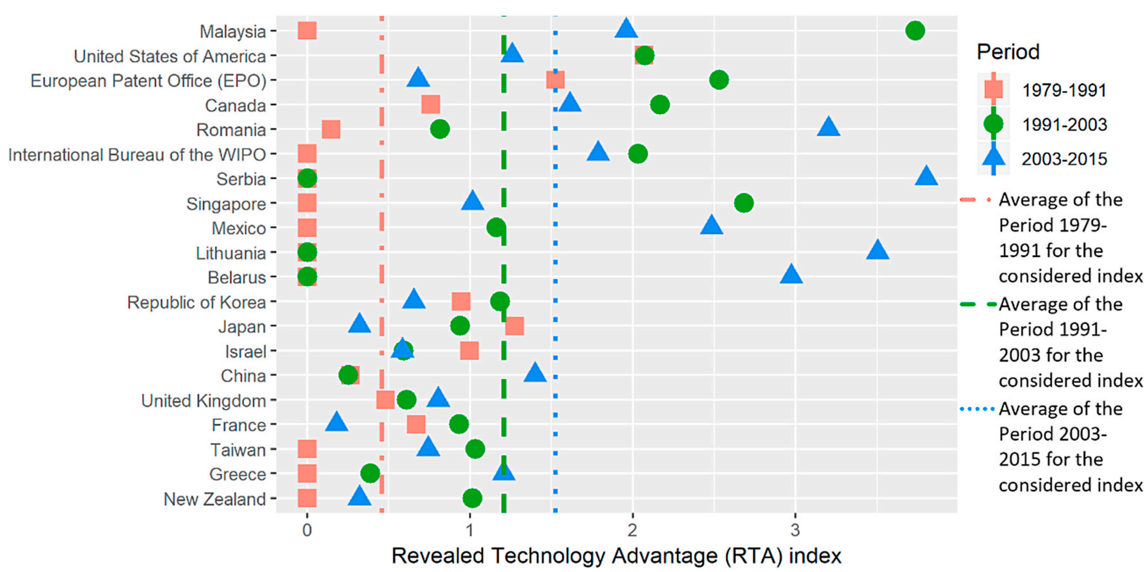

Fig. 3. Top 20 Patent Offices according to (and sorted by) the sum of their RTAs over three periods. Source: Authors' calculations based on data from PATSTAT 2017b.

WIPO, also had lower RTAs values in the last period; the EPO lost its positive specialisation, WIPO retained it.

5.2. National AI Breeding Grounds

Next, we operationalise the idea of Breeding Grounds, already introduced, in a national context. We use two basic assumptions for our indicator, later adding a third assumption. The first basic assumption is that the characteristic of a country as a National AI Breeding Ground depends on the number of AI-related priority inventions registered directly by this country's patent office. We assume that such patents represent an intention from inventors or companies to prioritise the exploitation of this country's market. The second assumption is that a country's RTA correlates with the extent to which it can be described as a NBG. A high RTA signals a specialisation into AI technologies in the country, which suggests the availability of relevant inputs, such as skilled labour, needed to exploit a technology. Combining both assumptions enables us to generate our indicator for NBGs, which is the product of the total number of patents related to AI with priority in a country in a period p, filed to the national patent office, and the country's RTA index for the given period p (see Eq. (3)).

$$Nat\ \ Breeding\ \ Ground_p = RTA\_Country_p \times \text{Priority Patents AI\_Country}_p \tag{3}$$

This indicator reduces the effects of high patenting numbers made by countries highly active in patent registration (either by the use of more flexible rules for the registration of patents by the country's application office, or because the country has a higher population, for example). At the same time, it allows us to deal with some of the known problems of using the RTA index, as elaborated in [9].

In addition to these two basic assumptions, we assume that PCT applications reflect more valuable inventions than non-PCT applications. Consequently, we calculate a weighted measure (see Eq. (4)).

$$\text{Weighted Patents\_AI\_Country}_p = \frac{1}{n}\sum_{t=1}^{n}\left(\left(\frac{\text{Non} - \text{PCT Patents AI\_Country}_{t,p}}{5}\right) + \left(\text{PCT Patents AI\_Country}_{t,p} \times 5\right)\right) \tag{4}$$

Here, '$\text{Non} - \text{PCT Patents AI\_Country}_{t,p}$' means the number of AI *priority* filings of 'A' type registered by the considered country at the year t for a given period p, while '$\text{PCT Patents AI\_Country}_{t,p}$' means the number of PCTs ('W' type) AI priority filings registered by the considered country during year t, which is in period p. With this measure, we calculated a slight variation from our previous Eq. (3), presented in Eq. (5).

$$Nat\ \ Breeding\ \ Ground\ \ Weighted\ \ Country_p = RTA\_Country_p \times Weighted\ \ Patents\_AI\_Country_p \tag{5}$$

The results of Eqs. (3) and (5) for the 15 patent offices[6] with the largest National AI Breeding Ground indicators are presented in Figs. 4 and 5, respectively. Again, the three vertical lines indicate the average of the presented values for each of the three periods considered. For better visualisation, the Logs10 of the calculated values are presented (both for the indices and for the averages).

Most remarkably, there is a substantial difference between China and the other patent offices, especially in the third period, as well as a large increase in both indicators for China between periods two and three, indicating that China became a NBG for AI more recently. At the same time, Japan and the US had been the main National AI Breeding Grounds in the 80s and the 90s; Japan declined markedly during the third period, while the US maintained its position as the second most relevant NBG.

Countries with high specialisation values but with an absolute low number of patents in AI, such as Serbia, Lithuania and Belarus (see Fig. 3), cannot be considered NBGs for AI (see Fig. 4). On the other hand, countries such as Germany, Russia and India come up in the ranking of NBGs. The increase in the average value of period three in relation to periods one and two is associated mainly with the increase in the values for China, India and Serbia in this period.

The high relevance of Asian countries is maintained (see Fig. 4), with three of the four main National AI Breeding Grounds pertaining to Asian patent offices, while the Eastern European patent offices from the European cluster are not so relevant (although Romania and Serbia appear at the bottom of Fig. 4). The patent offices from UK, Russia, Canada, Romania and Mexico can perhaps also be considered relevant National AI Breeding Grounds in the third period, but the relevance of Germany and France declines drastically.

The consideration of a higher weight for PCT patents reduces the mean of the patent offices' values (Fig. 5). China remains in the first place, but now with a lower value and a smaller difference to the US. There is an increase in the relevance of the Mexican patent office (regarding its position as a NBG). The consideration of the number of PCTs also favours Singapore, Belarus and Finland (comparing Figs. 4 and 5). On the other hand, when the number of PCTs is considered, the patent offices from Taiwan, the UK, Canada, Russian Federation, France and Romania lose positions as NBGs, as do the patent offices of Malaysia, India and Serbia (see Fig. 5).

5.3. International AI Breeding Grounds

The NBG indicators (weighted and not weighted) focus on the perspective of a single country. They deliberately neglect the registers of patents between countries. To take account of this international perspective, we developed an International Breeding Ground Index (weighted and not weighted). We make three assumptions for this index. First, countries with efficient IP protection and promising market potential for AI exploitation will attract a greater number of AI-inventions owned by inventors/companies from abroad. Second, countries with relevant AI development will seek the exploitation of their IP in promising foreign markets, reflected by their registration abroad of a higher number of national AI inventions. Finally, we assume that an IBG could be reflected by both types of international flows of patent registrations, which leads us to suggest a product rather than a sum for this index.

We introduce two new variables: '*AI patents coming from abroad_Country*$_{t,p}$' and '*AI patents going abroad_Country*$_{t,p}$', which, respectively, represent the number of AI priority filings from patent offices located in other countries registered in the country considered, and the number of priority filings registered by the patent office of this particular country that have also been patented in other countries, during a given year t and in a given period p. To identify which countries are International AI Breeding Grounds, we calculated the product of the

[6] Patent offices that cover more than one single country (such as the EPO, the International Bureau of the WIPO and the EAPO) are removed from Figs. 4 and 5, as they do not allow a particular country's market to be identified.

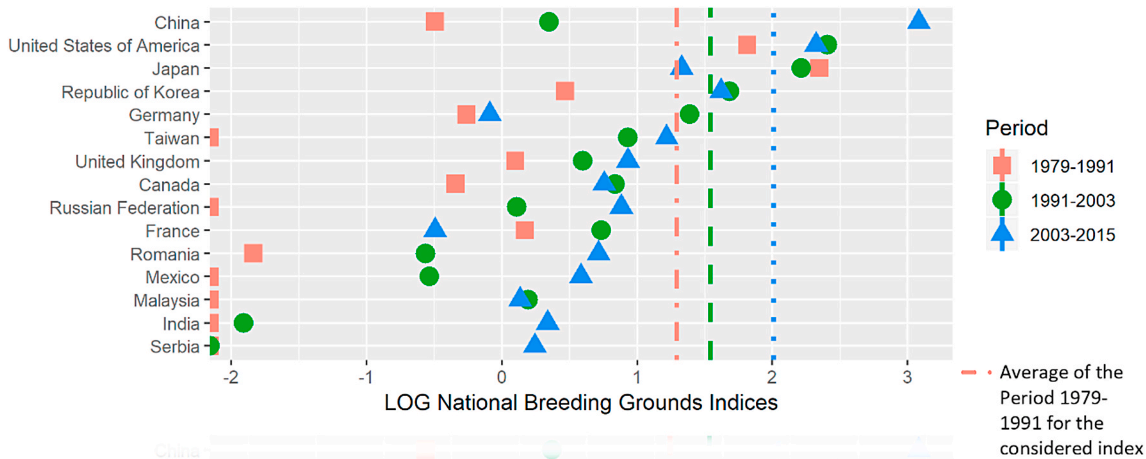

Fig. 4. Top 15 Patent offices which are considered National AI Breeding Grounds according to the Nat Breeding Ground_Country$_p$ indicator. Source: Authors' calculations based on data from PATSTAT 2017b.

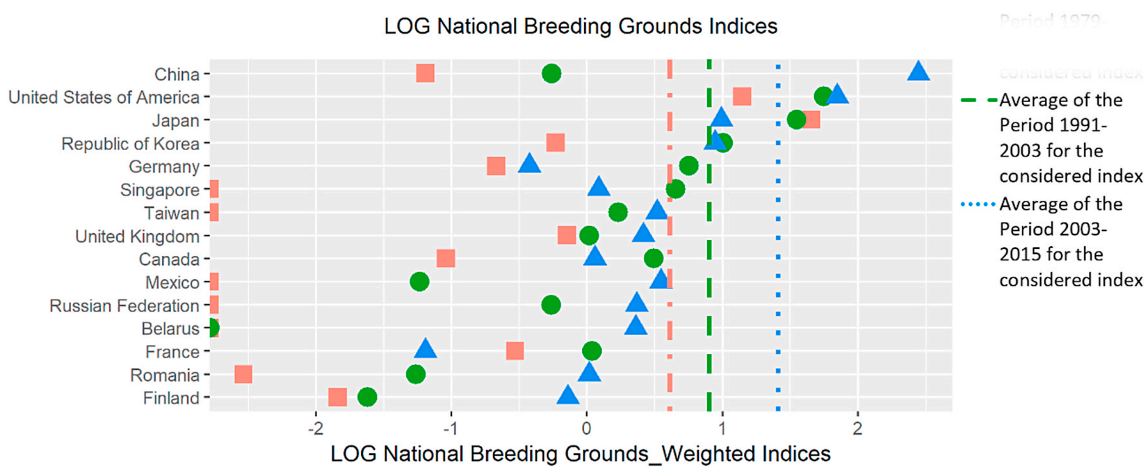

Fig. 5. Top 15 Patent offices which are considered National AI Breeding Grounds according to the Nat Breeding Ground_Weighted_Country$_p$ indicator. Source: Authors' calculations based on data from PATSTAT 2017b.

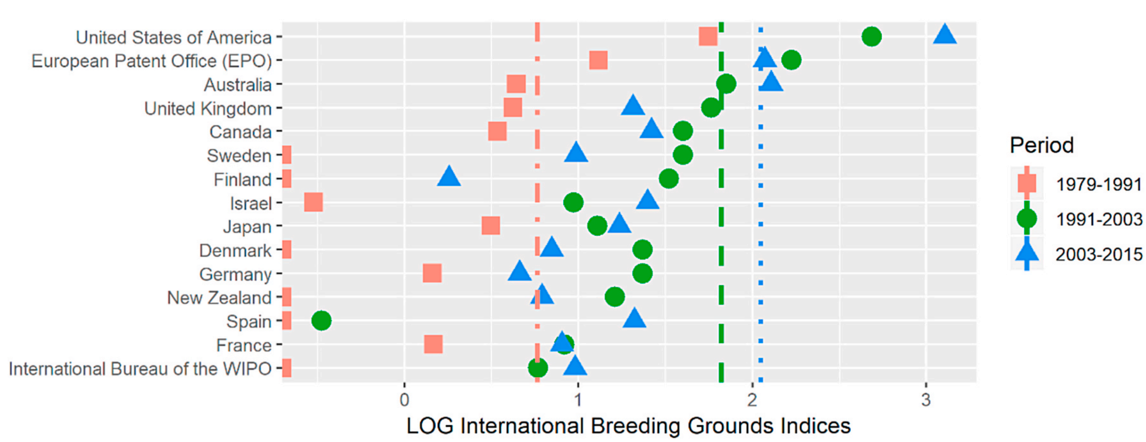

Fig. 6. Top 15 Patent Offices which are considered International Breeding Grounds according to the Int Breeding Ground_Country$_p$ indicator. Source: Authors' calculations based on data from PATSTAT 2017b.

two newly introduced variables and related it to two additional variables. First, we related it to the priority AI patents of this country (see Eq. (6)).

$$Int\ Breeding\ Grund_{country\_p} = \frac{1}{n}\sum_{t=1}^{n}\frac{AI\ patents\ coming\ from\ foreign\_country_{t,p} \times AI\ patents\ going\ to\ foreign\_country_{t,p}}{\text{Priority Patents AI\_Country}_{t,p}} \tag{6}$$

Second, we related it to the weighted indicator for NBG (see Eq. (7)) in order to find out the relationship between National and International Breeding Grounds while also considering the differences between PCT and non-PCT applications.

$$Int\ Breeding\ Ground_{Weighted\_country\_p} = \frac{1}{n}\sum_{t=1}^{n}\frac{AI\ patents\ coming\ from\ abroad\_Country_{t,p} \times AI\ patents\ going\ abroad\_Country_{t,p}}{Nat\ Greenhouse\_Weighted\_Country_{t,p}} \tag{7}$$

The results of these two measures are presented (see Figs. 6 and 7) again for the 15 patent offices[7] with the largest values, both using vertical lines to indicate the mean and the Logs10 of the calculated values for better visualisation.

We observe the US as a dominant International AI Breeding Ground, with an increase in both of its indicators over all periods (see Figs. 6 and 7). The US is followed by Australia and the EPO, which are characterised by different patterns: Australia has a similar pattern to the US, while the EPO shows a decrease during the latest period considered. This EPO pattern is also reflected in the trends of most European patent offices. Strikingly, while China is highly relevant as a National AI Breeding Ground, it has no relevance as an IBG. Furthermore, Sweden, Denmark and Spain, which are not considered National AI Breeding Grounds, now appear together with Finland as International AI Breeding Grounds. Moreover, Israel and Japan are shown to be constantly evolving as International AI Breeding Grounds.

In comparison with Fig. 6, the consideration of a higher value for PCTs in Fig. 7 increases the mean values. The consideration of PCTs also favours Australia, as well as Israel, Germany, New Zealand and France. Moreover, this consideration also shows that South Korea, India and Ireland have emerged as relevant International AI Breeding Grounds. On the other hand, the position of the EPO, Finland, Denmark and Spain is negatively affected, and WIPO, Spain and Finland disappear from the picture.

6. AI techniques: evolution, specialisations and Breeding Grounds

To analyse AI techniques we focused upon 12 techniques cited in more than 100 patents each. This time we made no distinction between the types of patents (PCTs and non-PCTs), thus looking only at the RTAs and at the (non-weighted) NBG and IBG values. As a general overview, the number of patents for each technique[8] is presented in Fig. 8.

We observe a large increase in the number of AI patents associated with the considered techniques (see Fig. 8). Far more patent applications related to Neural Networks, Expert systems and Fuzzy logic were registered in the second period than had been registered in the first. From the second to the third periods, however, the number of applications related to Fuzzy Logic systems decreased, while those associated with Machine Learning and Support Vector Machines further increased. By the end of third period, these two, together with Neural Networks, comprised the top three techniques. This movement at the top of the list is also seen farther down. In particular, Deep Learning had the most abrupt increase in this period; techniques related to Rule-based learning and Expert Systems had only a minimal growth in the third period, which might indicate a possible decline in the use of patents related to these techniques in recent applications when the drastic increase of AI patents in this period is considered.

Next we looked at the RTA, as well as the (non-weighted) NBG and IBG values, to identify the two leading Patent offices for each AI technique (see Table 3). This time, no distinction between the periods was made (thus, Period = 1 and t = 36), and only those AI techniques that at least 1,5% of the total sample were considered (thus we excluded Graphical Models, Classification and Regression Trees, Reinforced Learning, Logic Programming and Rule Learning techniques). Together, the seven analysed AI techniques comprise 95.7% of the sample.

It turns out that China leads again as a National AI Breeding Ground, being on the top in three out of seven AI techniques considered, followed by the US and Japan, which are the leaders of two and one AI techniques, respectively. When considering the International AI Breeding Ground index, the US stands out, being the leader of all AI techniques considered. Furthermore, the US can be considered the top NBG and IBG in Machine learning as well as in Supervised/Unsupervised Learning.

7. Discussion

The analysis of the presented data enables the identification of three perspectives on the evolution of AI. The first is related to the growth in the number of countries involved in AI in terms of specialisation. In this group, Asian and Eastern and Southern European countries stand out, with high levels of specialisation since the 90s and 2000s, respectively. Despite the fact that Eastern and Southeastern European countries

[7] For this indicator, international offices such as the EPO, WIPO and EAPO are not excluded. They can support International Breeding Grounds in the sense of attracting AI patents from abroad due to their economic potential as relevant AI markets. Similarly, patent offices without any priority filings are excluded.

[8] According to the same three periods and considering again the vertical lines for the mean values of each period as well as the Log 10 of the real values for better visualisation.

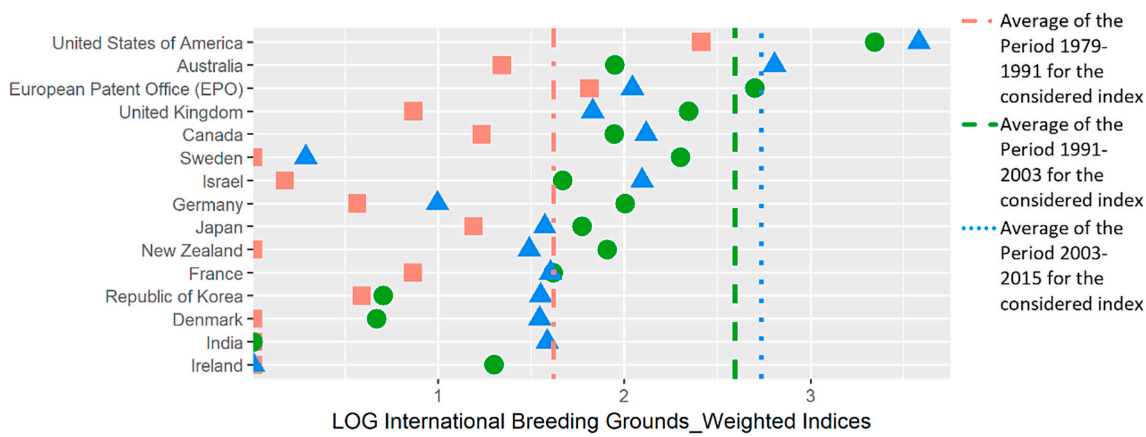

Fig. 7. Top 15 Patent Offices which are considered International Breeding Grounds according to the Int Breeding Ground_Weighted_Country$_p$ indicator. Source: Authors' calculations based on data from PATSTAT 2017b.

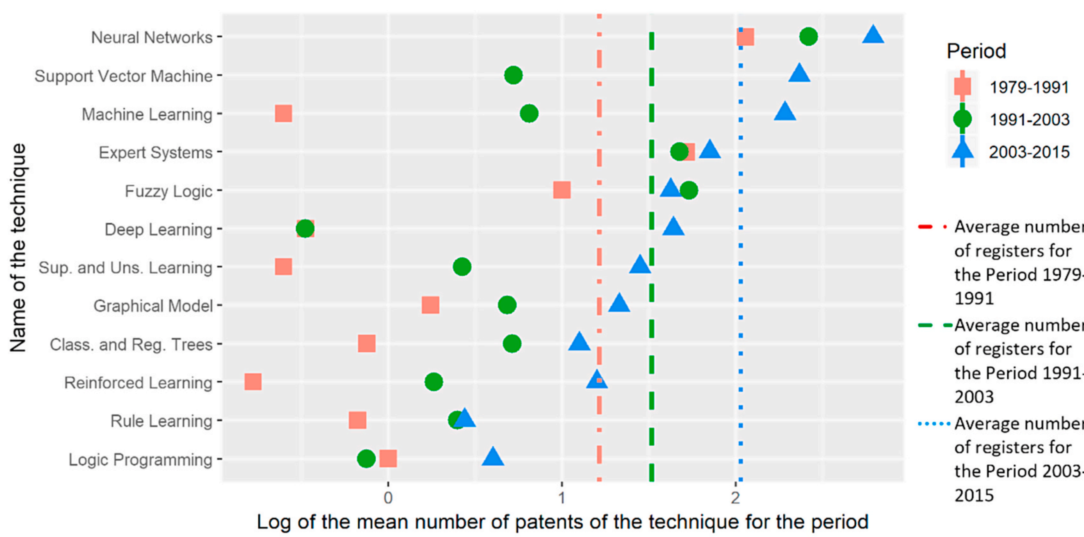

Fig. 8. Evolution of the patents of each of the AI-related techniques considered, sorted by the total number of patents for each technique. Source: Authors' calculations based on data from PATSTAT 2017b.

Table 3
National and International Breeding Ground leaders for each AI technique considered.

AI Technique	Position	RTA Index	National AI breeding ground leader	International AI breeding ground leader
Neural Networks	1st	Malaysia	China	US
	2nd	Serbia	Japan	EPO
Support Vector Machine	1st	India	China	US
	2nd	US	US	Australia
Machine Learning	1st	New Zealand	US	US
	2nd	US	China	Australia
Expert Systems	1st	Belarus	Japan	US
	2nd	Lithuania	US	EPO
Fuzzy Logic	1st	Mexico	US	US
	2nd	Romania	Germany	EPO
Deep Learning	1st	Japan	China	US
	2nd	Republic of Korea	Japan	Japan
Supervised and Unsupervised Learning	1st	Singapore	US	US
	2nd	Serbia	China	EPO

Source: PATSTAT 2017b. Authors' calculations.

specialised in AI later than other countries, a large proportion of their patents are related to AI techniques declining in global relevance. This is the case, for example, of the patent offices of Belarus and Lithuania in techniques related to Expert Systems, and of Romania in techniques related to Fuzzy Logic.

Furthermore, results from analysing NBG and IBG values suggest that Japan and most West European countries have lost their early vanguard status, while China and the US have increased their leadership. China is leading in a cluster of National AI Breeding Grounds, whereas the US is not only a significant NBG for AI but is also the leading International AI Breeding Ground for each and every analysed technique.

This signals a major structural difference in the international patterns for IP protection in AI after the 1990s: some countries have focused on developing their domestic markets and have been less interested in the exploitation of foreign markets, while others have developed AI in an international context. Asian patent offices are by and large in the first group (China, India, Malaysia, Republic of Korea, Taiwan), as are the Russian Federation, Mexico, Romania and Serbia; whereas developed countries from the western hemisphere are the dominant IBGs.

Finally, our findings suggest substantial changes in the relevance of various AI techniques over time, as already prominently documented

[16]. In particular, we confirm that the use of mathematical models (like fuzzy logic), as well as knowledge-based models (like Expert Systems) is decreasing, while Biological and Machine Learning Models (such as neural networks, supervised and unsupervised machine learning, deep learning and Support Vector Machines) are increasing their relevance as AI techniques. Our investigation adds to these findings by considering the international competition in leadership as well as the above-explained structural differences between the US and China at the level of AI techniques. Both the US and China have top positions in terms of NBGs in the AI techniques considered, but only the US is a top NBG and IBG in AI techniques. This indicates that the US leads not only in these rapidly growing AI techniques at home but that this happens also in an international context of IP protection. In contrast, China seems to sustain its leading position by focusing on IP protection of its domestic market, almost in isolation from the international context.

Although, the increasing relevance of China in the global arena for AI development cannot be ignored, the structural features identified could be related to several underlying factors. On a very general level, China is characterised by a state-capitalistic approach: AI has become a high political priority in the last decade [16,37]. It has been documented [12, 14] that in China, universities rather than corporate actors account for a large majority of AI-related patenting. This applies also to [16], which found that 98% of all registrations at SIPO in the sample originate from Chinese universities. On the other hand, the US is more associated with a market-driven 'Silicon Valley approach' to AI, which is more open and internationally connected. Furthermore, Chinese leadership as the top National AI Breeding Ground is reduced when considering PCT-applications, which signal that China-produced IP is less relevant for protection in international markets [12] points out further evidence suggesting that Chinese universities' activities are reflected in their patenting behaviour but are not based on inventions marketable enough for international protection. Thus, it is very likely that the output in China-related AI patenting is connected to the incentive structure for Chinese researchers at universities. It is also likely that the extent of commercialisation of AI IP by universities and corporate actors differs, which might help to explain why China scores so low as an IBG, since the prospect of commercialisation motivates corporate actors to seek to extend IP protection in foreign jurisdictions. However [12], further highlights that Chinese MNC subsidiaries are indeed aiming to protect their local IP on the Chinese market, but that they are still not generating innovations relevant enough to file on international lead offices. Moreover, our results could indicate that foreign corporate actors do not register AI-related IP with SIPO, which might be related to institutional barriers and/or limitations in terms of local enforcement [12].

8. Conclusions

This paper analyses trends and structural differences in patenting patterns in AI-related technologies. We propose two novel patent-based indicators to differentiate structural differences in the patterns of IP protection observed for AI across countries. We considered (i) to what extent countries specialise in AI and are relevant markets for corresponding IP protection ('National Breeding Ground'); and (ii) to what extent countries attract AI from abroad for IP protection and extend IP protection of their own AI to foreign markets ('International Breeding Ground'). We demonstrate that NBGs and IBGs overlap only to a limited extent. Primarily, China and the US can be characterised as dominant NBGs. Australia, selected European countries, but primarily the U.S., are major IBGs. We conclude that China promotes AI development with an almost-exclusive focus on IP protection in the domestic market, whereas the US sustains its AI progress in an international context, too. This might indicate a considerable bifurcation in structural patterns of IP protection in global AI development. We discussed possible explanations related to the institutional particularities of the Chinese National Innovation System.

This paper contributes to the broader debate by introducing and operationalising the concepts of NBGs and IBGs. The proposed approach, in general, can be used as a reference for further patent mining and technology innovation analysis of other technical or scientific fields. However, we acknowledge limitations in our approach, which include the data source, the method, and the indicators themselves. First, we based our analyses exclusively on the patent process. However, actors might use other means to disclose their inventions (e.g., by defensive publication in scientific journals, pre-print servers or platforms); attention exclusively on patents might miss valuable innovations in the area of AI. Nevertheless, we were interested in longer-term trends and dynamics, and it is here that patents offer a suitable source of internationally standardised information available in a longer time series. Second, we used priority filings as a proxy for the first market on which companies, other organisations and inventors aim to protect an invention. By doing this, we can indicate the market impact as well as the markets impacted by a given technology. However, this approach neglects the development aspects of AI, which could be captured by the location of the inventors. Furthermore, we use patent applications rather than granted patents, which introduces potentially non-relevant IP. Some applicants are not granted patents, as many as seven years might be required before the patent office makes a decision and, most problematic, both the proportion of applications granted and the processing time vary from office to office. Third, we try to account for the quality of patent application by considering PCT registers. Yet, there exist more comprehensive quality measures for patents, such as citations, renewal rates or high-impact inventions. Incorporating these kinds of data could help to reduce the 'noise' that surely appears in the comparative analysis presented here. Fourth, we pay limited attention to variations in the particular characteristics of patentability between patent offices (e.g., highlighted in [23]), which might also introduce a 'quality bias' into our dataset. Finally, we restricted our keyword search to title and abstract. This could be improved by considering the claims in the whole patent document. The challenge would be to differentiate dependent and independent claims in the identification strategy, but if this challenge could be overcome, this approach could not only improve the identification strategy, but also reduce the potential risk of 'fashionable labelling' trends that discourage the use of keyword-based searches.

CRediT authorship contribution statement

Matheus Eduardo Leusin: Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Jutta Günther:** Conceptualization, Resources, Writing - review & editing, Supervision. **Björn Jindra:** Conceptualization, Methodology, Investigation, Validation, Writing - review & editing, Supervision. **Martin G. Moehrle:** Conceptualization, Methodology, Validation, Project administration, Investigation, Writing - original draft.

Declaration of competing interest

The authors declare no competing interests.

APPENDIX A. QUERY FOR IDENTIFYING AI APPLICATION IDS

Select appln_id from tls202_appln_title
where appln_title like '%Artificial intelligence%' OR appln_title like '%machine learn%' OR appln_title like '%Probabilistic reason%' OR appln_title

like ‘%Fuzzy logic%’ OR appln_title like ‘%Logic Programming%’ OR appln_title like ‘%Ontology engineer%’ OR appln_title like ‘%pervised learn%’ OR appln_title like ‘%reinforced learn%’ OR appln_title like ‘%task learn%’ OR appln_title like ‘%neural network%’ OR appln_title like ‘%deep learn%’ OR appln_title like ‘%expert system%’ OR appln_title like ‘%support vector machin%’ OR appln_title like ‘%description logistic%’ OR appln_title like ‘%classification tree%’ OR appln_title like ‘%regression tree%’ OR appln_title like ‘%logical learn%’ OR appln_title like ‘%relational learn%’ OR appln_title like ‘%probabilistic graphical model%’ OR appln_title like ‘%rule learn%’ OR appln_title like ‘%instance-based learn%’ OR appln_title like ‘%latent represent%’ OR appln_title like ‘%bio-inspired approach%’ OR appln_title like ‘%machine intelligen%’ OR appln_title like ‘%probability logic %’ OR appln_title like ‘%probabilistic logic%’ OR appln_title like ‘%reinforcement learn%’ OR appln_title like ‘%multitask learn%’ OR appln_title like ‘%Decision tree learn%’ OR appln_title like ‘%support vector network%’ OR appln_title like ‘%deep structured learn%’ OR appln_title like ‘%hierarchical learn%’ OR appln_title like ‘%graphical model%’ OR appln_title like ‘%structured probabilistic model%’ OR appln_title like ‘%Rule induction %’ OR appln_title like ‘%memory-based learn%’ OR appln_title like ‘%bio-inspired comput%’ OR appln_title like ‘%biologically inspired comput%’

UNION

select appln_id from tls203_appln_abstr

where appln_abstract like ‘%Artificial intelligence%’ OR appln_abstract like ‘%machine learn%’ OR appln_abstract like ‘%Probabilistic reason%’ OR appln_abstract like ‘%Fuzzy logic%’ OR appln_abstract like ‘%Logic Programming%’ OR appln_abstract like ‘%Ontology engineer%’ OR appln_abstract like ‘%pervised learn%’ OR appln_abstract like ‘%reinforced learn%’ OR appln_abstract like ‘%task learn%’ OR appln_abstract like ‘%neural network%’ OR appln_abstract like ‘%deep learn%’ OR appln_abstract like ‘%expert system%’ OR appln_abstract like ‘%support vector machin%’ OR appln_abstract like ‘%description logistic%’ OR appln_abstract like ‘%classification tree%’ OR appln_abstract like ‘%regression tree%’ OR appln_abstract like ‘%logical learn%’ OR appln_abstract like ‘%relational learn%’ OR appln_abstract like ‘%probabilistic graphical model%’ OR appln_abstract like ‘%rule learn%’ OR appln_abstract like ‘%instance-based learn%’ OR appln_abstract like ‘%latent represent%’ OR appln_abstract like ‘% bio-inspired approach%’ OR appln_abstract like ‘%machine intelligen%’ OR appln_abstract like ‘%probability logic%’ OR appln_abstract like ‘% probabilistic logic%’ OR appln_abstract like ‘%reinforcement learn%’ OR appln_abstract like ‘%multitask learn%’ OR appln_abstract like ‘%Decision tree learn%’ OR appln_abstract like ‘%support vector network%’ OR appln_abstract like ‘%deep structured learn%’ OR appln_abstract like ‘%hierarchical learn%’ OR appln_abstract like ‘%graphical model%’ OR appln_abstract like ‘%structured probabilistic model%’ OR appln_abstract like ‘% Rule induction%’ OR appln_abstract like ‘%memory-based learn%’ OR appln_abstract like ‘%bio-inspired comput%’ OR appln_abstract like ‘%biologically inspired comput%’

APPENDIX B. IPC-search strategy adopted in papers used for comparison

	Fujii & Managi, 2018 (Query 2)	Tseng & Ting, 2013 (Query 3)
IPC	codes	G06 N 3/00, G06 N 3/02, G06 N 3/04, G06 N 3/06, G06 N 3/063, G06 N 3/067, G06 N 3/08, G06 N 3/10, G06 N 3/12, G06 N 5/00, G06 N 5/02, G06 N 5/04, G06 N 7/00, G06 N 7/02, G06 N 7/04, G06 N 7/06, G06 N 7/08, G06 N 99/00
G05B 13/ 02,	G06E 1/00, G06E 3/00, G06F 9/44, G06F 15/00, G06F 15/18, G06F 17/00, G06F 17/20, G06G 7/00, G06J 1/00, G06 N 3/00, G06 N 3/02, G06 N 3/04, G06 N 3/ 08, G06 N 3/10, G06 N 3/12, G06 N 5/00, G06 N 5/02, G06 N 5/04, G06 N 7/00, G06 N 7/02, G06 N 7/04, G06 N 7/06, G06 N 7/08, G06 N 99/00	

APPENDIX C. Variables considered and their definitions, according to the authors

Variable	Meaning	Source
Global number of AI patents$_{t,p}$	Total number of patents related to AI globally in year t in a period p	Patstat 2017
Global number of patents$_{t,p}$	Total number of patents related to AI globally in year t in a period p	Patstat 2017
Priority patents_country$_{t,\ p}$	Total number of patents with priority in a country in year t in a period p	Patstat 2017
Priority patents AI_country$_{t,\ p}$	Total number of AI patents with priority in a country in year t in a period p	Patstat 2017
No PCT patents AI_country$_{t,\ p}$	Total number of patents related to AI with priority in a country in year t in a period p (‘A’ type patents in PATSTAT), filed to the national patent office	Patstat 2017
PCT patents AI_country$_{t,\ p}$	Total number of patents related to AI with priority in a country in year t in a period p (‘W’ type patents in PATSTAT), filed to WIPO	Patstat 2017
Weighted patents AI_country$_p$	Indicator related to the total number of AI patents with priority in a country in a period p, but with the distinction of taking different weights for PCT and non-PCT applications.	Authors calculations
AI patents coming from abroad_country$_{t,\ p}$	Total number of patents related to AI with priority in a different country in year t in a period p	Patstat 2017
AI patents going abroad_country$_{t,\ p}$	Total number of patents related to AI going abroad with priority in a country in year t in a period p	Patstat 2017
RTA_country$_p$	Revealed technology advantage of a country in period p	Authors calculations
Nat breeding ground_country$_p$	Indicator for National Breeding Ground in a country in period p	Authors calculations
Nat breeding ground_weighted_country$_p$	Indicator for National Breeding Ground in a country in period p, but with the distinction of taking different weights for PCT and non-PCT applications.	Authors calculations
Int breeding ground country$_p$	Indicator for International Breeding Ground in a country in period p	Authors calculations
Int breeding ground_weighted_country$_p$	Indicator for International Breeding Ground in a country in period p, but with the distinction of taking different weights for PCT and n non-PCT applications.	Authors calculations
t	Index for years in a period p (ranging from 1 to 12 for each period p)	According to definition
p	Index for periods (ranging from 1 to 3)	According to definition
n	Number of years in a period p (12 years in our case)	According to definition

References

[1] X. Li, H. Jiang, Artificial intelligence technology and engineering applications, Appl. Comput. Electromagn. Soc. J. 32 (5) (2017).

[2] J.M.-G.J. Klinger, K. Stathoulopoulos, Deep Learning, Deep Change? Mapping the Development of the Artificial Intelligence General Purpose Technology, 2018.

[3] N.J. Nilsson, The Quest for Artificial Intelligence, Cambridge University Press, 2009.

[4] N.F. Bassis, F. Armellini, Systems of innovation and innovation ecosystems: a literature review in search of complementarities, J. Evol. Econ. 28 (5) (2018) 1053–1080.

[5] M.P. Hekkert, R.A. Suurs, S.O. Negro, S. Kuhlmann, R.E. Smits, Functions of innovation systems: a new approach for analysing technological change, Technol. Forecast. Soc. Change 74 (4) (2007) 413–432.

[6] P. Sharma, R. Tripathi, Patent citation: a technique for measuring the knowledge flow of information and innovation, World Patent Inf. 51 (2017) 31–42.

[7] M.J. Mansourzadeh, B. Shahmoradi, H. Dehdarirad, E. Janavi, A note on using revealed comparative advantages in scientometrics studies, Scientometrics 121 (1) (2019) 595–599.

[8] B. Balassa, Trade liberalisation and "revealed" comparative advantage 1, Manch. Sch. 33 (2) (1965) 99–123.

[9] L. Soete, The impact of technological innovation on international trade patterns: the evidence reconsidered, Res. Pol. 16 (2–4) (1987) 101–130.

[10] M.A. Weresa, Technological competitiveness of the EU member states in the era of the fourth industrial revolution, Econ. Bus. Rev. 5 (3) (2019).

[11] K. Menzel, L. Maicher, A novel method for retrieving specialisation profiles–The case of patent agent firms, World Patent Inf. 51 (2017) 46–56.

[12] H. Kroll, Exploring the validity of patent applications as an indicator of Chinese competitiveness and market structure, World Patent Inf. 33 (1) (2011) 23–33.

[13] Ö. Gürerk, B. Irlenbusch, B. Rockenbach, The competitive advantage of sanctioning institutions, Science 312 (5770) (2006) 108–111.

[14] WIPO, WIPO Technology Trends 2019: Artificial Intelligence, World Intellectual Property Organization, Geneva, 2019.

[15] C.-Y. Tseng, P.-H. Ting, Patent analysis for technology development of artificial intelligence: a country-level comparative study, Innovation 15 (4) (2013) 463–475.

[16] H. Fujii, S. Managi, Trends and priority shifts in artificial intelligence technology invention: a global patent analysis, Econ. Anal. Pol. 58 (2018) 60–69.

[17] L. Huang, W. Miao, Y. Zhang, H. Yu, K. Wang, Patent network analysis for identifying technological evolution: a case study of China's artificial intelligence technologies, management of engineering and technology (PICMET), in: 2017 Portland International Conference on, IEEE, 2017, pp. 1–9.

[18] J. Niu, W. Tang, F. Xu, X. Zhou, Y. Song, Global research on artificial intelligence from 1990–2014: spatially-explicit bibliometric analysis, ISPRS Int. J. Geo-Inf. 5 (5) (2016) 66.

[19] L. Aristodemou, F. Tietze, The state-of-the-art on Intellectual Property Analytics (IPA): a literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (IP) data, World Patent Inf. 55 (2018) 37–51.

[20] A. Annoni, P. Benczur, P. Bertoldi, B. Delipetrev, G. De Prato, C. Feijoo, E. F. Macias, E.G. Gutierrez, M.I. Portela, H. Junklewitz, Artificial Intelligence: A European Perspective, Joint Research Centre (Seville site), 2018.

[21] C.A. Keisner, J. Raffo, S. Wunsch-Vincent, Breakthrough Technologies–Robotics, Innovation and Intellectual Property, WIPO2015.

[22] I.M. Cockburn, R. Henderson, S. Stern, The Impact of Artificial Intelligence on Innovation, National Bureau of Economic Research, 2018.

[23] K. Meguro, Y. Osabe, Lost in patent classification, World Patent Inf. 57 (2019) 70–76.

[24] M. Lupu, Artificial intelligence and intellectual property (editorial), World Patent Inf. 53 (2018) A1–A3.

[25] E. Gabrilovich, S. Markovitch, Overcoming the Brittleness Bottleneck Using Wikipedia: Enhancing Text Categorization with Encyclopedic Knowledge, AAAI, 2006, pp. 1301–1306.

[26] P. Sharma, R. Tripathi, R. Tripathi, Finding similar patents through semantic query expansion, Procedia Computer Science 54 (2015) 390–395.

[27] EPO, Data Catalog – PATSTAT EP Register – 2018, Autumn Edition, 2018.

[28] A. Radauer, C. Rosemberg, O. Cassagneau-Francis, H. Goddar, C. Haarmann, Study on the Economic Impact of the Utility Model Legislation in Selected Member States: Final Report. A Study Tendered by the European Commission–DG Internal Market and Services in 2013, MARKT/2013/065, 2015.

[29] G. De Rassenfosse, H. Dernis, G. Boedt, An introduction to the Patstat database with example queries, Aust. Econ. Rev. 47 (3) (2014) 395–408.

[30] WIPO, The PCT now has 152 Contracting States. https://www.wipo.int/pct/en/pct_contracting_states.html, 2019.

[31] J. Lapenne, Patent cooperation treaty (PCT), J. Pat. & Trademark Off. Soc'y 92 (2010) 192.

[32] G. De Rassenfosse, H. Dernis, D. Guellec, L. Picci, B.v.P. de la Potterie, The worldwide count of priority patents: a new indicator of inventive activity, Res. Pol. 42 (3) (2013) 720–737.

[33] G. Thoma, Composite value index of patent indicators: factor analysis combining bibliographic and survey datasets, World Patent Inf. 38 (2014) 19–26.

[34] N. Van Zeebroeck, B. Van Pottelsberghe de la Potterie, Filing strategies and patent value, Econ. Innovat. N. Technol. 20 (6) (2011) 539–561.

[35] J.P. Martino, Technological Forecasting for Decision Making, McGraw-Hill, Inc., 1993.

[36] P. Zuniga, D. Guellec, H. Dernis, M. Khan, T. Okazaki, C. Webb, OECD Patent Statistics Manual, OECD Publications, Francia, 2009.

[37] D. Harhoff, S. Heumann, N. Jentzsch, P. Lorenz, Outline for a German Strategy for Artificial Intelligence, 2018.

Matheus E. Leusin is a PhD candidate in the working group of Prof. Dr. Jutta Günther at the University of Bremen and is currently researching the worldwide economic competition for the leading role in artificial intelligence as part of his dissertation. His experience involves the analysis of Innovation Systems based on bibliometric and patent data and the use of modeling and simulation methods for prospective analysis. During his master's degree, completed in 2018, Matheus analysed the diffusion of wind technology in Brazil using a Technological Innovation Systems approach. Based on this analysis, incentive policies for wind diffusion in Brazil were proposed, which were further evaluated through scenario analysis involving simulation. Matheus also worked as a guest researcher at the Bremer Institut für Produktion und Logistik GmbH (BIBA), conducting research on the development of Cyber-Physical Systems.

Prof. Dr. Jutta Günther studied economics and social science at the Universities of Oldenburg, Osnabrück and Albany/New York from 1994 to 1999. In 2002 she did her doctorate at the University of Osnabrück in the field of social science at the chair of economic theory. From 2002 to 2014 she worked at the Leibniz Institute for Economic Research Halle (IWH) as scientist. She held different leading positions there, for example as head of department for structural change from 2008 to 2013 or as member of the executive board from 2011 to 2013. She habilitated at the faculty for economic science of the Friedrich-Schiller-University of Jena in 2014. In April 2014 Jutta Günther took the professorship of economics, especially innovation and structural change at the University of Bremen. Her research priorities are innovation processes, structural change and economic aspects of system transformation.

Prof. Dr. Björn Jindra completed a PhD in Science and Technology Policy (University of Sussex-SPRU) and a Habilitation in Economics (University of Bremen), and currently holds an Associate Professor position at the Department of International Economics, Government and Business of the Copenhagen Business School. His research is at the intersection of international business, innovation and development studies. Previous work focused primarily on the role of multinational firms for the technological development of emerging economies. His current work deals with the internationalization of R&D by European multinational firms as well as the global race for leadership in artificial intelligence technologies.

Prof. Dr. Martin G. Moehrle holds the Chair of Business Administration with a focus on Innovation and Competence Transfer and Director of the Institute for Project Management and Innovation (IPMI) at the University of Bremen. Previously, he headed the Chair of General Business Administration and Specialisation in Planning and Innovation Management at the Brandenburg Technical University Cottbus. His research interests lie in the various facets of innovation management, including the market-driven design of novel products, processes and services, the methodical induction of inventions with TRIZ, the prognosis of technological and market developments and the management of intellectual property rights. He is an industrial engineer of the TH Darmstadt with a specialisation in electrical engineering.