January 2023 · Pegah Maham, Stefan Heumann,
Wiebke Denkena, Laurenz Hemmen,
Anna Semenova

AI Talent Flows in Germany

Empirical study of the career paths of AI doctoral students at German universities

Stiftung
Neue
Verantwortung

Think Tank at the Intersection of Technology and Society

Acknowledgements

Our team at Stiftung Neue Verantwortung (SNV) deserves our gratitude for the support throughout the research and publication process. Within SNV, we would like to thank in particular Lisa Koeritz, Shannon Reitmeir, Luisa Seeling, Alina Siebert, Anna-Katharina Meßmer and Martin Degeling.

We would also like to thank Oluwabimpe Abiodun Lawal for the support with the creation of the dataset, Remco Zwetsloot who shared his valuable experience regarding talent flow analyses with us, Bijan Soltani for his advice on data infrastructure, Elke Müller for the translation of the German paper and Alexis Close for helpful final revisions.

The paper belongs to the SNV programme 'Data Science', which is supported by Stiftung Mercator.

Executive Summary

Discussions of the geopolitical implications of artificial intelligence (AI) often narrowly spotlight the competition between the USA and China. Assessments and comparisons of their respective strengths and weaknesses with regard to the availability of data, skilled engineers and scientists, and supercomputing infrastructure receive a lot of attention. Control over these key prerequisites determines who can best leverage AI for their own interests and geopolitical aspirations.

While the USA and China feature prominently in the discourse on the geopolitics of AI, the contributions of other regions and countries and their efforts to position themselves in the context of global cooperation and competition for the development of AI are often not part of the picture. Although the European Union (EU) is currently discussed as an important player in AI regulation, efforts to better understand the strengths and weaknesses of Europe's AI ecosystem remain limited.

This pilot study attempts to shed light on Europe's AI ecosystem by offering empirical insights into the mobility of AI researchers in Europe. The ability to attract and retain AI talent is a crucial indicator of the strengths and weaknesses of national AI ecosystems. It is well known that the strength of the US AI ecosystem rests on its ability to attract and retain the best global talents.

In this data brief, we analyse the pool of top AI talents for the EU's most populated country: Germany. Our aim is to better understand how Germany is integrated into global AI talent flows. Where does Germany draw talent from? Where do young researchers go after completing their doctorates in Germany? What can the data potentially tell us about the strengths and weaknesses of the German AI ecosystem? Our analysis focuses on a self-built dataset of the career paths of PhD students supervised by the most prominent AI professors in Germany.

Our analysis offers insights into talent flows in and out of Germany. Half of the PhD students in our sample received their undergraduate degrees at foreign universities. While EU countries were found to play a much less important role as countries of origin than anticipated, China, India and Iran are more important than expected. Most PhD students remain in Germany for at least a few years after graduation. However, international doctorates, in particular, tend to leave the country after graduation. Our data show that while the USA, the UK and Switzerland hardly send any PhD students to Germany, they are important destinations for AI talent from German universities. Global tech companies that offer high salaries and research budgets are the most important employers of talents from Germany in these countries.

Our pilot study shows the potential of data-based investigations on the mobility of AI talents in Germany. We welcome further analyses of talent flows and other key factors of success in the European context. They will enable the development of measures to support Europe's AI ecosystem in a targeted and evidence-based manner. Furthermore, such empirical insights help keep track of whether political objectives are achieved and what their actual impacts are (e.g., attracting top researchers from foreign countries or strengthening the exchange between EU member states).

Table of Contents

Introduction

Artificial intelligence (AI) is widely regarded as one of the key technologies of the 21st century. As a so-called general purpose technology like electricity or the steam engine, AI is predicted to enhance and potentially revolutionize many different applications across all sectors of our economy, ranging from medical diagnostics and fraud detection to autonomous driving and smart energy grids. As a dual-use technology, AI is also expected to change the way wars are fought, with applications ranging from intelligence collection and analysis to autonomous weapons systems. Given how far and wide-ranging the potential use cases are, it is no wonder that the question of which country has an edge in the development and deployment of AI technologies has become central to discussions on international power and geopolitics.

Those concerned with the geopolitical implications of AI have devoted significant attention to the competition between the USA and China for technological supremacy. If there is anything like a race for AI development, Europe and other regions of the world often do not seem to be part of it as real competitors – they only play a role as battlegrounds on which the two 'AI superpowers' confront each other.[1] However, it is important to recognize that the development of AI occurs within networks that often involve collaboration across national borders. The focus on the USA and China ignores the contributions of other regions and countries and their efforts to position themselves in the context of global cooperation and competition for the development of AI.

This paper is an effort to begin filling the gap that has been created by the narrow spotlight placed on China and the USA. European countries and the European Union (EU) have formulated ambitious strategies and policies to position themselves globally, not only as regulators but also as innovators of AI technologies. These policy documents receive global attention, and Brussels is widely seen as a 'regulatory superpower'.[2] However, there has been little empirical analysis of Europe's AI capabilities. Aside from the availability of data and supercomputing infrastructure, we believe that the ability to attract and retain AI talent is a crucial indicator of the strengths and weaknesses of national AI ecosystems, as it is well known that the strength of the US AI ecosystem rests on its ability to attract and retain the best global talents.[3]

1 Lee, K.-F. (2018). *AI superpowers: China, Silicon Valley, and the new world order.* Houghton Mifflin Harcourt.

2 Bradford, A. (2020). The Brussels effect. *Northwestern University Law Review*, 107(1), 25–66. https://doi.org/10.1093/oso/9780190088583.003.0003

3 Zwetsloot, R., Dunham, J., Arnold, Z., & Huang, T. (2019). *Keeping top AI talent in the United States.* Center for Security and Emerging Technology. https://doi.org/10.51593/20190007

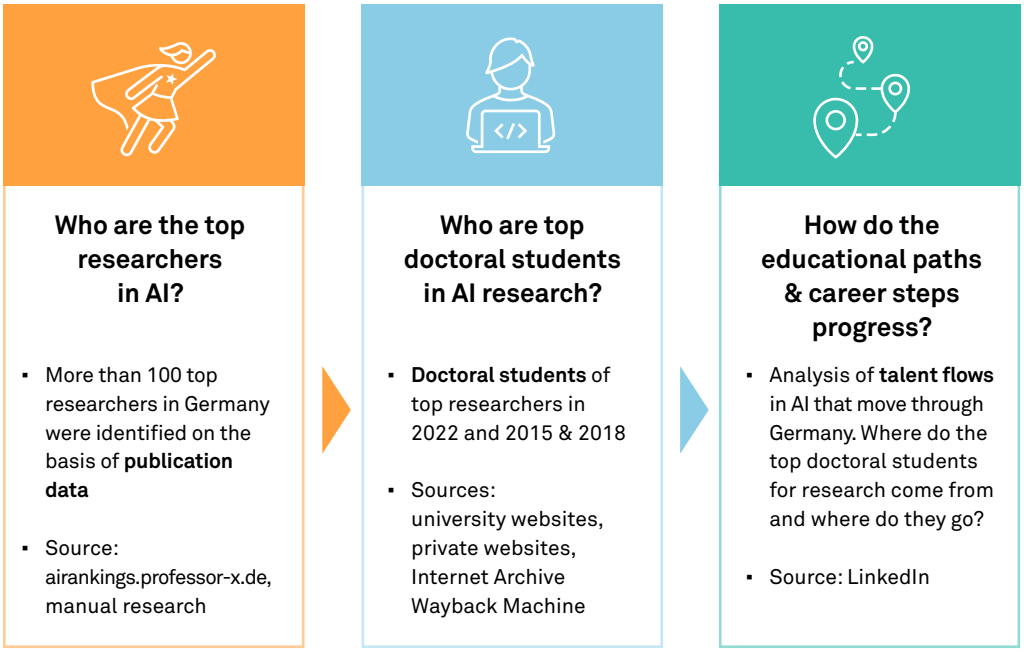
In this data brief, we analyse the pool of top AI talents for the EU's most populated country: Germany. Our aim is to better understand how Germany is integrated into global AI talent flows. Where does Germany draw talent from? Where do young researchers go after completing their doctorates in Germany? What can the data potentially tell us about the strengths and weaknesses of the German AI ecosystem? Our analysis focuses on a self-built dataset of the career paths of PhD students supervised by the most prominent AI professors in Germany. The analysis of this dataset allowed us to draw some preliminary conclusions to the questions posed above. We explain our approach and methodology in Chapter 1.

Our analysis provides insights into talent flows in and out of Germany. Half of all PhD students in our sample received their undergraduate degrees at universities abroad. EU countries were found to play a much less important role as countries of origin than anticipated, while China, India and Iran are much more important as sources of AI talent than expected. Most PhD students remain in Germany, at least for the next few years after graduation. However, our data also confirm that while the USA, the UK and Switzerland hardly send any doctoral students to Germany, they are important destinations for AI talents from German universities. Global tech companies that offer high salaries and research budgets are the most important employers for talents from Germany in these countries.

Methodical approach

Our analysis focuses on the top young scientists in Germany in the various sub-disciplines in the field of 'artificial intelligence', such as Natural Language Processing, Computer Vision or Machine Learning. In this context, we focus on the group of doctoral students who are supervised by internationally recognized top researchers as the competition for these individuals is particularly intense. In order to obtain information about educational and career paths of this group, we have collected data in three stages (see figure 1). First, we have created a list of top researchers in AI in Germany on the basis of publication data. Secondly, starting from this list, we have identified current and former doctoral students of these top researchers (referred to as top doctoral students). Finally, with regard to these doctoral students, we collected data on their educational and career paths.[4]

Figure 1



The figure shows the main steps for data collection and processing on which our analysis is based. First, we identified leading AI researchers using publication data. Then, we researched the names of their doctoral students and collected resume data via LinkedIn.

We will explain the individual stages below. Annex 1 contains a more detailed description of our data sources and methodology.

4 For this analysis, we have processed personal data for scientific research purposes. If you suspect that our analysis is based on data relating to you and you wish to exercise your right of objection, please follow the instructions in our data protection notice on our website.

Who conducts top research in AI in Germany?

In order to identify top researchers in AI in Germany, we have drawn on publication data from 13 renowned international conferences that reflect different sub-disciplines of the field (see Annex 1, section A). For the purposes of our analysis, individuals were considered as top researchers who are currently working in Germany and who have presented and published at such conferences at least twelve times in the period from 2012 to 2022. We were able to identify a total of 111 top researchers in Germany.[5] 90 of them are supervising doctoral students at German universities.

Who are the top PhD students in AI?

Based on our list of top researchers in AI in Germany, we have created two datasets regarding top PhD students that we will briefly explain below; please find further details on the creation of these datasets in Annex 1, section B.

Dataset 1: Current doctoral students (2022)
By means of manual research, we have created a list of doctoral students who are currently pursuing their PhD and are supervised by at least one top researcher in this process. We were able to collect the names of the doctoral students of almost all the top researchers through the websites of the corresponding academic chair or research lab at the respective universities or the personal websites of the top researchers. In this way, we identified a total of 893 doctoral students of top researchers in AI.

Dataset 2: Former doctoral students (2015 & 2018)
In order to obtain a sample of individuals who have earned a doctoral degree in AI in the past and were supervised by a top researcher in this process, we used archived websites that we accessed by means of the Internet Archive Wayback Machine. Based on archived web pages from the years 2015 and 2018, we were able to identify a total of 806 names of former doctoral students who were supervised by a top researcher in the past. Dataset 1 contains 142 of these names.[6] The two datasets together contain the names of 1557 current and former top doctoral students in AI in Germany.

5 Annex 1, section A contains a more detailed description of these 111 top researchers.

6 In most of the cases, this overlap occurs because the doctoral studies of the respective individuals extend over a lengthy period. Figure 9 in Annex 2, section B shows the exact distribution of the start and termination years of the doctorates of the doctoral students in our sample.

How do the educational paths and career steps of the top PhD students progress?

In order to collect data on the educational paths and first career steps of the top PhD students identified, we have drawn on information that these individuals themselves have published in their LinkedIn profiles. LinkedIn is a social network, which is used to maintain professional contacts. It is widely used in Germany, also among academics. By means of manual research, we were able to identity the LinkedIn profiles of 58 per cent of the current and former doctoral students contained in our datasets, which could be clearly associated with the relevant individuals by means of self-reported information on the doctorate in the profile. In Annex 4, we explain assessments and approaches to control potential biases in our findings that are attributable to self-selection effects (selection bias).

From the total of 898 LinkedIn profiles, we obtained information such as the university attended as well as previous jobs and employers. More than 80 per cent of the profiles contain information on the respective bachelor's degree. 95 per cent of the profiles contain data on the current career stage (doctorate in progress or current profession). In 710 of the cases, i.e. 79 per cent of the profiles, we were able to obtain both of these data points. According to their own information, approximately half of the doctoral students in our sample have already completed their doctorate.[7] Annex 2, section A contains a tabular overview of the data basis of our analysis. Please find in Annex 1, section C a detailed description of how we analyzed the information obtained from the profiles.

7 With the help of the German National Library, we have randomly checked whether former doctoral students have successfully completed their doctorate. As we have found a published doctoral thesis for the vast majority of them, we assume that the number of uncompleted doctoral projects is negligible and we will treat, in the following, all former doctoral students as 'graduates' of a doctoral programme.

Findings

Countries of origin of top doctoral students in AI at German universities

Our data show that Germany is attractive internationally as a location for doctoral studies in AI. More than half (53 per cent) of the PhD students identified earned their first university degree not in Germany but in foreign countries (see figure 2). Although we do not have information on their citizenship or overall educational background, it can be assumed that the individuals in our sample who did not earn their first university degree in Germany are mostly international students (foreign students). We conclude this e.g. from the fact that most of them have studied in countries that are not typically countries to which German citizens go for their first-degree studies.[8]

For the sake of comparison: Across all subjects and disciplines, the proportion of international doctoral students at German universities amounted to approximately 12 per cent in the winter semester of 2020/21.[9] Thus, there are significantly more international doctoral students in AI, in relative terms, than at German universities in general.[10] However, compared to the USA, German top locations for doctorates in AI attract fewer talents from abroad. In the USA, the proportion of international doctoral students in AI related disciplines amounts to 64 per cent on average.[11]

8 46 per cent of the non-German bachelor's degrees were earned in China, India, Iran or Russia. However, most of the individuals with German citizenship who leave Germany for their first-degree studies go to the Netherlands and Austria. See: Statistisches Bundesamt [German Federal Bureau of Statistics] (2021). *Deutsche Studierende im Ausland – Ergebnisse des Berichtsjahres 2019* [German students in foreign countries – findings in the reporting year 2019].

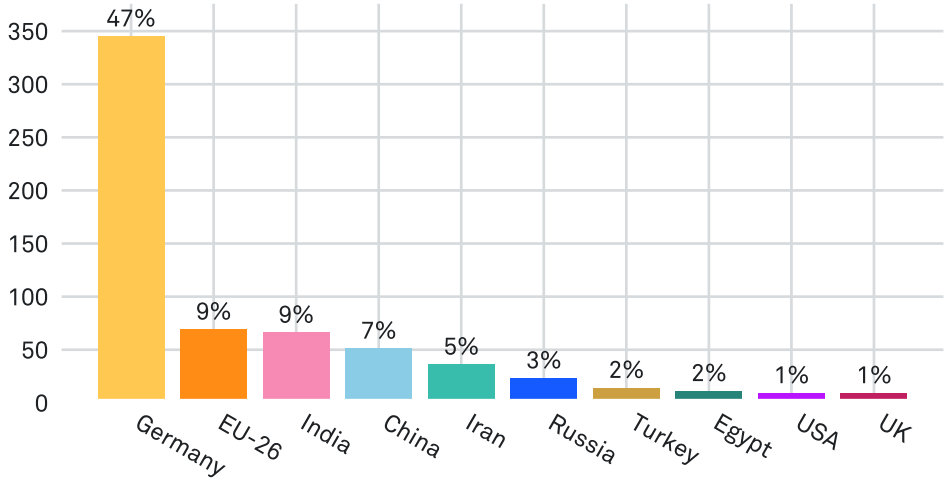
9 DAAD [German Academic Exchange Service] & DZHW [German Center for Research on Higher Education and Science] (Eds.) (2022). *Wissenschaft weltoffen kompakt: Facts and Figures on the International Nature of Studies and Research in Germany and Worldwide*.

10 In that context, the proportion of international doctoral students differs, depending on the location in Germany (see figure 10 in Annex 2, section C).

11 Zwetsloot, R., Dunham, J., Arnold, Z., & Huang, T. (2019). *Keeping Top AI Talent in the United States*. Center for Security and Emerging Technology.

Figure 2

Where did doctoral students in AI in Germany earn their first university degree?

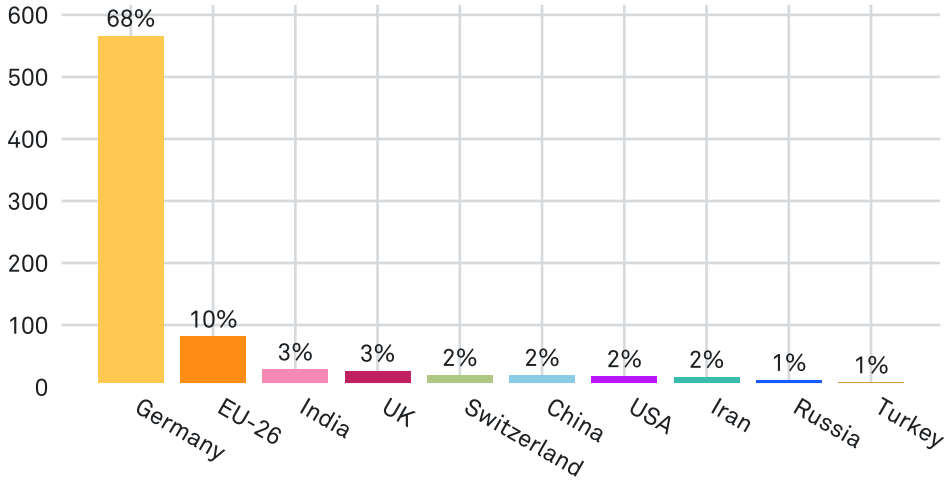


The figure shows the countries in which most of the current and former doctoral students in our sample, about which we have been able to collect data on their bachelor's degree, earned their first university degree (n=727). The y-axis shows the absolute number of individuals. In total, 87 per cent of the top doctoral students in AI at German universities earned their first university degree in one of the countries indicated above. More than half of them earned their bachelor's degree in a country other than Germany. The other countries most represented are India, China and Iran comprising up to 9 per cent each. The category EU-26 includes all current EU member states other than Germany, and together account for 9 per cent.

Many of the international top doctoral students in AI in Germany came to Germany not for their doctorate, but already for their master's studies (see figure 3). 68 per cent of the doctoral students completed their master's degree at a German university (compared to 47 per cent of bachelor's degrees earned in Germany). 9 per cent previously studied at a university in another EU country. In our sample, relatively few doctoral students completed their master's degree in a non-European country (22 per cent).

Figure 3

Where did doctoral students in AI in Germany get their master's degree?



The figure shows the most common countries where current and former doctoral students earned a master's degree. This sample is based on students for which we were able to collect information regarding their master's degree (n=833). The y-axis shows the absolute number of individuals. In total, 95 per cent of the top doctoral students in the AI field at German universities earned their master's degree in one of the countries indicated above. Almost 80 per cent of the master's degrees were earned in Germany and the other EU member states (EU-26). Other countries account for a maximum of 3 per cent – significantly fewer than in the case of bachelor's degrees.

Destination countries after graduation

On average, 63 per cent of the 396 former doctoral students in our sample who already completed their doctorate and documented their current employer on LinkedIn are still working in Germany (see figure 4). In this context, individuals in our sample completed their doctorate on average three years ago (see Annex 2, section A).

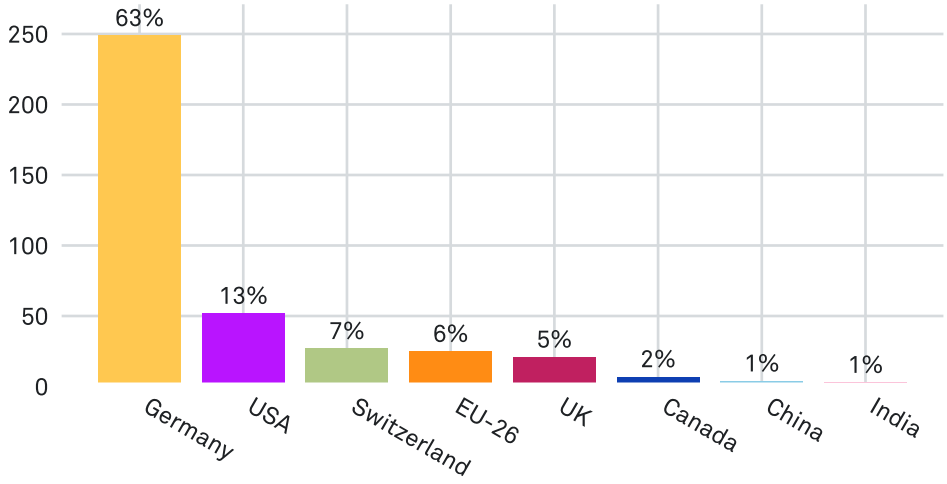
Of the individuals in our sample who earned their first university degree outside of Germany (mainly international doctoral students), fewer individuals than on average stay in Germany after completing their doctorate. Only 54 per cent of them are currently working in Germany. Compared to the USA, where 82 per cent of the top international graduates in AI are still working in the country even five years after their doctorate,[12] the proportion of international top graduates who stay in Germany after completing their doctoral degree is significantly lower.

12 Zwetsloot, R., Dunham, J., Arnold, Z., & Huang, T. (2019). *Keeping Top AI Talent in the United States*. Center for Security and Emerging Technology.

The main countries of destination for people leaving Germany are the USA, Switzerland and the United Kingdom. Together with Germany, this small group of countries attracts the majority (88 per cent) of the people with doctorates completed in Germany who had come to Germany from many different regions of the world. Switzerland alone attracts as many graduates from Germany as all the other 26 EU member states taken together. Although many of the former doctoral students completed their bachelor's degrees in India, China or Iran, few of them return to these countries in the first few years after their doctorate (see figure 4).

Figure 4

Where are former doctoral students in AI in Germany working today?



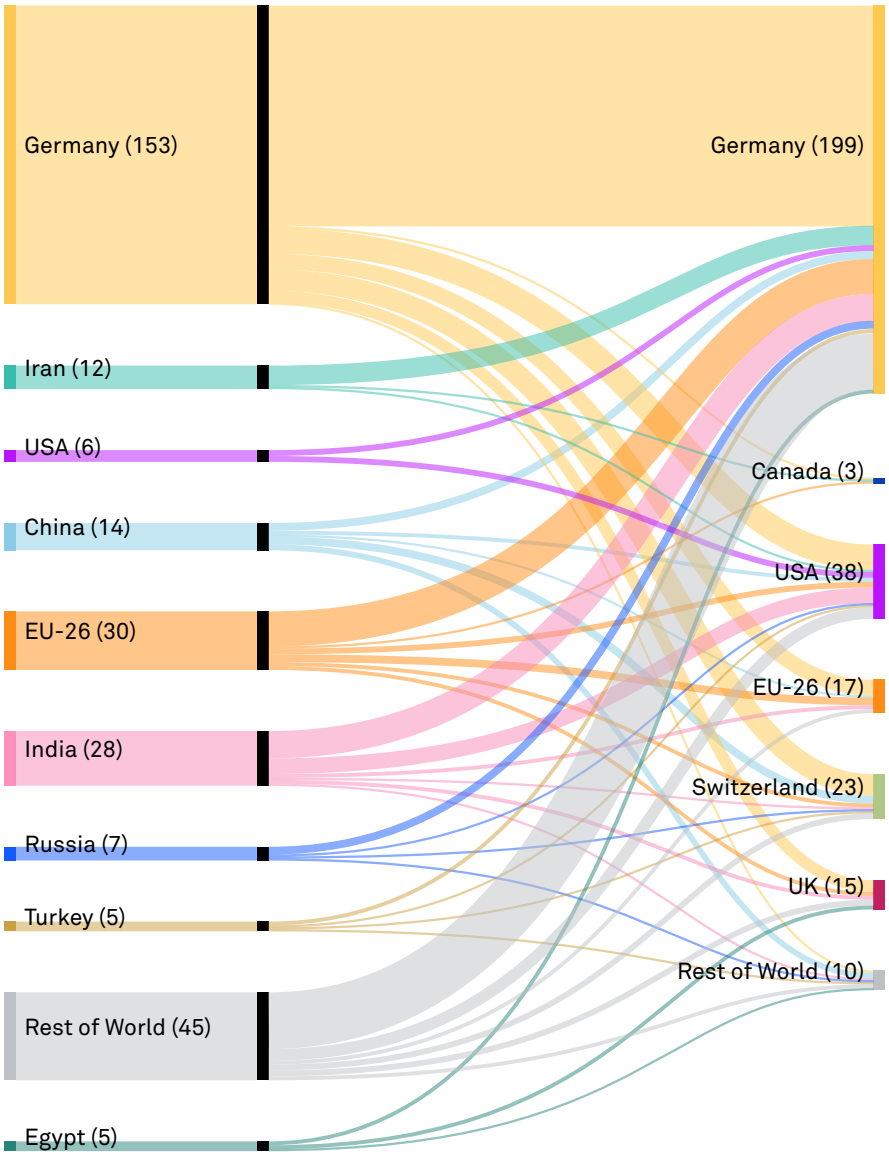
The figure shows the most common countries where former doctoral students in our sample are currently employed according to their LinkedIn profiles (n=396). The y-axis shows the absolute number of individuals. More than 98 per cent of the former top doctoral students in AI at German universities are now working in one of the countries indicated above. Approximately 63 per cent are working in Germany. About 13 per cent are currently working in the USA, followed by approximately 7 per cent working in Switzerland and about 6 per cent working in the EU (excluding Germany). China, India and Iran play no significant role here, although many people completed their bachelor's degrees in these countries.

Our data shows that after completing a doctoral degree in Germany, many of the top young scientists move from the East to the West. Numerous people who completed their bachelor's degree in India, China or Iran and came to Germany for their doctorate remain in Germany or move on to another European country, the UK, Switzerland or the USA (see figure 5). The few individuals who came to Germany for doctoral studies from these popular destination countries return to these same countries after their doctorates if they do not stay in Germany.

The post-doctoral migration strongly depends on the country of origin (see figure 5 and Annex 3, section C). Approximately one-third of the individuals who completed their first university degree in Germany leave Germany after completing their doctorate. About 50 per cent of the people who earned their bachelor's degree in India and came to Germany for their doctorate leave Germany after earning their doctoral degree. A particularly large number of people with a Chinese bachelor's degree leave Germany after earning their doctorate (about 70 per cent). In contrast, many individuals who have previously studied in Iran stay in Germany: approximately 80 per cent of them are currently working with a German employer.

The USA – the country with the highest inflow in our sample – seems to be an attractive destination country in particular for PhD graduates at German universities who earned their first university degree in India. While about half of them currently have a job in Germany, about a quarter of them now work in the USA.

Figure 5 Mobility of AI PhD Graduates in Germany



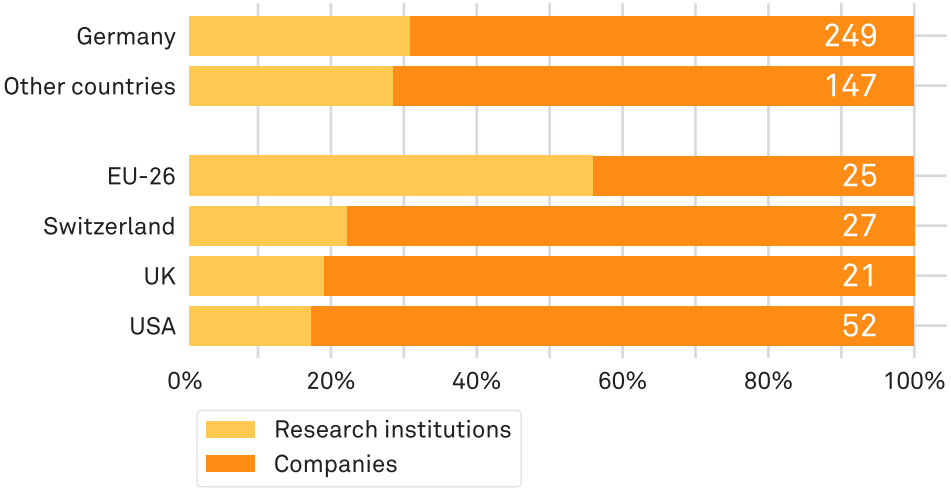
The figure gives an overview of the mobility of former doctoral students in our sample for whom we have obtained information on their bachelor's degree, the completion of their doctorate and their current job (n=305). In each case, the width of the horizontal bars corresponds to the proportion of people who earned their bachelor's degree in the respective country (left) and work in the respective country after having completed their doctorate in Germany (right). After the country names, the number of individuals is noted in brackets. The category 'Rest of World' (to the left and the right) covers all countries not already indicated in the respective list. Annex 3, section C shows the underlying data in tabular form.

Current employers of former AI doctoral students

Of the graduates of doctoral programmes who remained in Germany after completing their degrees, about 70 per cent work in the private sector (see figure 6). Among the graduates who left Germany after completing their doctorate, the proportion of private-sector employees is similarly high. Among those who accepted a job in the USA, the UK or Switzerland after completing their doctoral dissertation, a particularly high number are employed by a company (about 80 per cent in each case). The majority of the 25 graduates in our sample who moved to other EU countries work in research institutions (56 per cent).

Figure 6

Do former doctoral students work in research or in the private sector today?



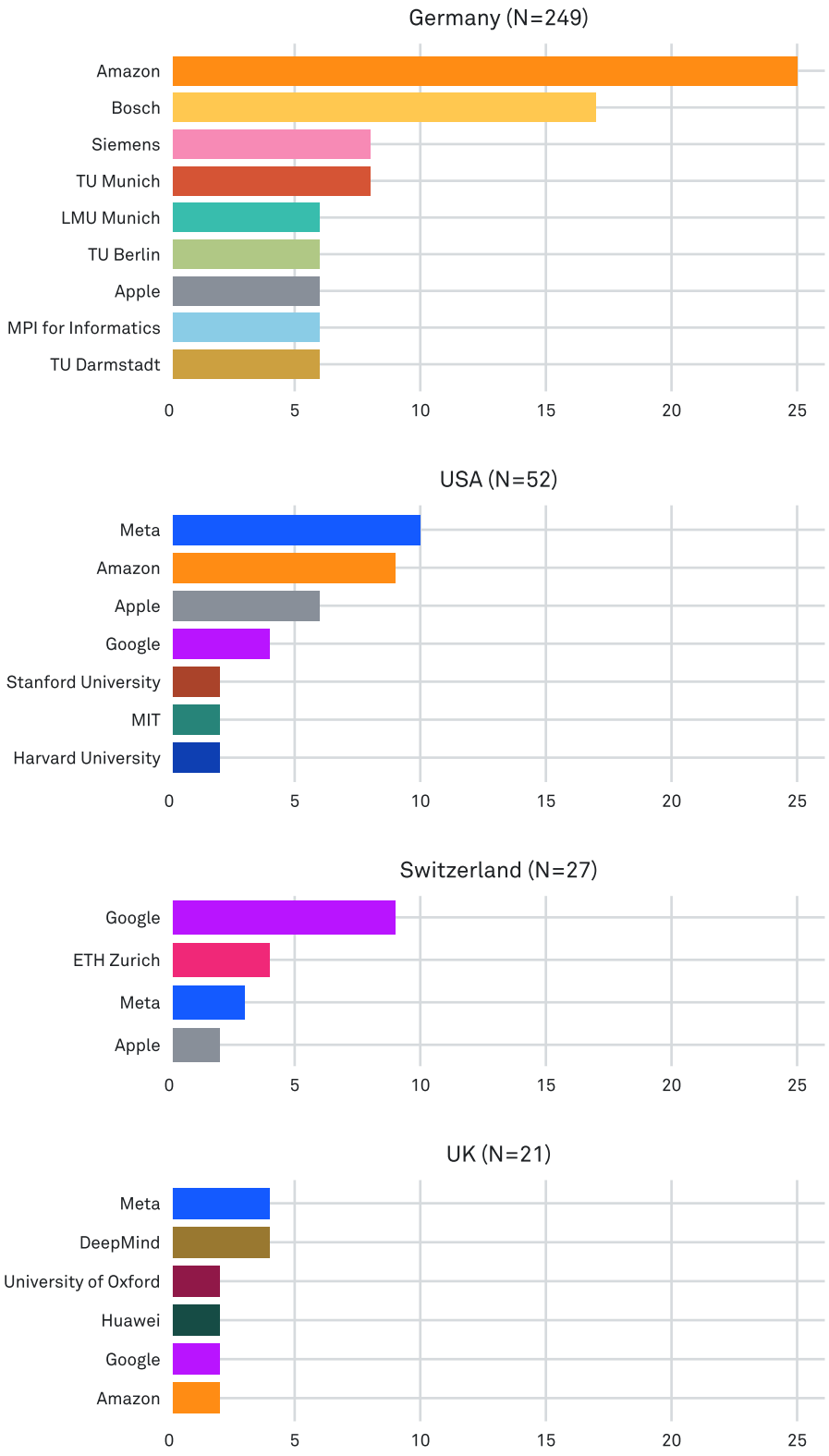
The figure shows the proportions of former doctoral students in our sample who have indicated their job on LinkedIn and are currently employed by a research institution or a company (n=396). The numbers on the bars indicate the number of individuals in our sample who are currently working in the respective country. The proportion of employees in research institutions and companies varies from country to country. Of the top PhD students who have completed their doctorate at a German university and are currently working in different EU countries, 56 per cent are working at a research institution and 44 per cent in the private sector. Significantly more of the individuals who now have a job in the USA, the UK or Switzerland work in the private sector (about 80 per cent).

Many of the individuals who left Germany after completing their doctoral dissertation now work for a big US tech company such as Alphabet (Google & DeepMind), Meta or Amazon (see figure 7). Renowned universities such as Oxford, Stanford or the ETH Zürich [Swiss Federal Institute of Technology Zurich] also attract graduates from Germany. In Germany, Amazon is a popular employer, closely followed by Bosch and the Technical University of Munich.

The number of former doctoral students who now work with certain employers differs between the destination countries. In the USA, Switzerland and the UK, most of them are employed by the four most popular employers (58 per cent, 72 per cent and 57 per cent in each case). In Germany, however, only a total of 28 per cent of the former doctoral students are employed with the four most popular employers.

Figure 7

Which are the top employers, in the most popular destination countries, for former AI doctoral students in Germany?



The figure shows the employers, in the most popular destination countries, that most frequently employ former doctoral students in our sample who have included current occupation information in their LinkedIn profile. The numbers in parentheses indicate the total number of individuals in our sample who currently work in the respective country. In this case, the sample for each destination country is larger than the sample we used to create figure 5, since we have also taken into account the individuals for whom we were unable to obtain information regarding their first university degree. The data shows: many of those who have completed their doctorates in Germany are now working at branches of large US tech companies in Germany and other countries.

Conclusions and outlook

This pilot study provides valuable first insights into the mobility of early-career AI researchers in Germany. Our data show that German universities strongly benefit from the influx of talents from other countries. More than half of the doctoral students in our sample earned bachelor's degrees outside of Germany. This is a much higher rate of international students than German universities attract on average. Considering that the EU has its own AI strategy and promotes collaboration among AI researchers across member states through numerous funding programmes, the small proportion of doctoral students in Germany who come from neighbouring European countries is surprising, especially compared to the many young researchers who come to Germany from India, China and Iran.

While Germany benefits from the influx of many international PhD students, our data also shows that international doctorates are particularly likely to leave Germany after graduation. Among the international doctoral students in our sample, 46 per cent left Germany after graduation in favour of employment abroad. Among those who received their undergraduate degrees and PhDs in Germany, only 31 per cent left the country after graduation. On average, the individuals in our sample finished their doctoral studies three years ago. Therefore, it is possible that the number of graduates who leave Germany rises even further as time passes.

In general, the USA, United Kingdom and Switzerland are popular destinations for PhD graduates in AI from Germany. About 25 per cent of the doctoral students in our sample moved to one of these three countries after earning their doctorates at a German institution. The list of employers in these countries is dominated by global tech companies with large AI budgets. Our data thus confirm what has been discussed under the catchphrase 'brain drain' in Germany: German universities and research institutions play a vital role in the development of the AI talent pool from which big tech companies recruit their staff. While offering valuable first empirical insights into the competition for talent in the European AI ecosystem in the example of Germany, our analysis does not permit us to draw conclusions about the reasons behind the migration patterns we observed. To understand the motivations of young AI researchers in pursuing their PhDs in Germany or in leaving the country after graduation, more qualitative research is needed (e.g., via interviews and surveys). Such insights could help policymakers develop targeted and evidence-based policies to better promote AI in Europe and track whether political objectives are actually being achieved.

Although more research is needed, our findings indicate that in the German context, there is room for improvement when it comes to the retention rates of young AI researchers, especially international doctoral students who earned their first university degree abroad. Many of these people leave Germany shortly after graduation. Furthermore, the attractiveness and competitiveness of German doctoral programs vis-à-vis the offerings of neighbouring European countries seems to be an issue. These are potential leverage points that the German government should further investigate and address.

Beyond the national scope, we believe that it is also important to increase efforts to understand the European AI ecosystem as a whole. Therefore, we welcome further case studies of AI mobility in other European national contexts. Taken together, such research would form a solid empirical basis for informing strategic policy decisions at the European level. To maintain its foothold and secure its position among global competitors in the long term, Europe needs to develop targeted and evidence-based measures to efficiently and impactfully support its own ecosystem. This will only be possible if policymakers have a deep understanding of the dynamics at play and know the leverage points to change them.

Annex 1: Addendum to the methodical approach

A Identifying the top researchers

In order to identify top researchers in Germany, we have relied on publication data of 13 international AI conferences (see table 1), which can be retrieved by means of the tool AI Author Rankings (airankings.professor-x.de). When selecting the conferences, we oriented ourselves towards the AI category in CSRankings,[13] a widely used international ranking of institutions in the field of computer sciences. In these conferences, for example, the sub-disciplines Natural Language Processing and Computer Vision, which have been heavily influenced by deep learning methods during the past 10 years, are strongly represented.

Table 1

AAAI	AAAI Conference on Artificial Intelligence
IJCAI	International Joint Conference on Artificial Intelligence
NeurIPS	Conference on Neural Information Processing Systems
ICML	International Conference on Machine Learning
KDD KDD	The International Conference on Knowledge Discovery and Data Mining
WWW	World Wide Web Conference
SIGIR	ACM SIGIR Conference on Research and Development in Information Retrieval
ACL	Annual Meeting of the Association for Computational Linguistics
EMNLP	Empirical Methods in Natural Language Processing

Table: International AI conferences considered in our analysis to identify top researchers by means of publication data.

13 In contrast to AI Author Rankings, CSRankings turned out to be unsuitable for our analysis, as the data relies on manual input and not all relevant individuals and institutions are included in the ranking.

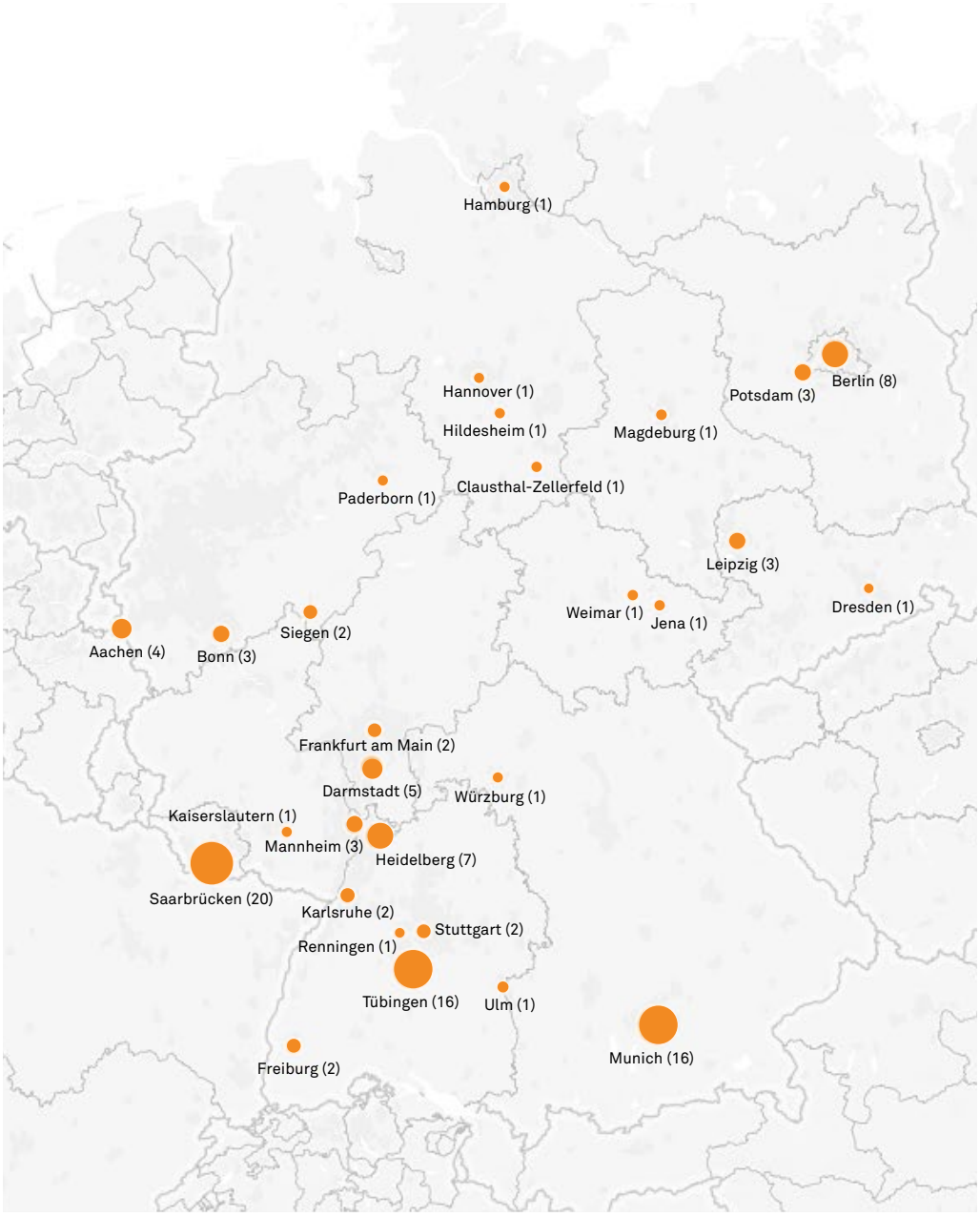
Using AI author rankings, we have identified 3796 individuals worldwide who contributed more than 12 publications between January 2012 and August 2022 at the conferences listed in table 1. These individuals were included in our analysis as top global researchers. We have deliberately not used the filter 'Germany' (provided by AI Author Rankings) in order to identify top researchers in Germany within this group, as the underlying data on the institutional affiliation of the researchers contained in the Digital Bibliography & Library Project (DBLP) database is often outdated or no information is available at all. Instead, we have manually researched the current affiliation of all 3796 top researchers worldwide.

In this way – through manual research of the institutional affiliation of the world's top researchers – we have identified 111 individuals who are currently (as of August 2022) working in Germany and who published at least 12 papers at one of the 13 international AI conferences between 2012 and August 2022. We were able to identify doctoral students connected to 90 out of these 111 top researchers. Most of the other 21 researchers do not supervise any doctoral students, as they had only very recently finished their doctorates or are employed in the private sector.

The German city in which most of the top researchers live is Saarbrücken (20), followed by Munich and Tübingen (16 each), Berlin (8) and Heidelberg (7) (see figure 8). In these cities, the top researchers are spread across several institutions, with many having double or even triple affiliations.

The vast majority of top researchers (86) holds titles such as Full Professor, Honorary Professor, Junior Professor or Associate Professor and works, at least partially, at an academic institution or at a research institute. Only 8 individuals appear to be employed exclusively in the private sector, 4 of them at one of the Amazon locations in Germany.

Figure 8 Where in Germany do top AI researchers work?



The map shows the places of employment of the 111 top AI researchers working in Germany who published, between January 2012 and August 2022, more than 12 times at one of the 13 international AI conferences listed in Annex 1. After the names of the cities, the number of people employed in the respective city is noted in brackets. Saarbrücken, Munich and Tübingen are the cities where most of the top researchers are based.

B Identifying current and former doctoral students

The lists with names of current and former doctoral students of the top researchers are based on manual research using online search engines and the Internet Archive Wayback Machine. They are obtained through the websites of the corresponding academic chair or research lab at the respective universities or the personal websites of the top researchers. In most of the cases, the doctoral students could be clearly identified in the years 2015, 2018 and 2022. In a few cases, no archived version of the corresponding page was available for the year 2015 or 2018. In such cases, we used archived pages of the years 2014, 2016 or 2017, if available.

Due to our approach (identifying doctoral students of researchers in AI with many publications), we could have included people into our dataset who do not work in AI but are or were supervised by one of the top AI researchers on another topic. In order to avoid such cases, we randomly checked the research focus of the doctoral candidates. With a few exceptions, the thematic reference to AI was clear.

A few top researchers are part of larger research groups with several professors who supervise doctoral students. In some cases it was not possible to find out which doctoral students are or were supervised by which professors. In such cases, all doctoral students of the research group were included in our dataset (if their research focuses on AI, which we ensured by means of manual online research).

In our analysis, we explicitly focus on doctoral students who will receive or have already received their doctoral degree at German universities. PhD students who work at a German institution in the scope of a cooperation such as the 'Cambridge-Tübingen AI Fellowships' and are supervised by a top researcher identified by us, but will earn their degree at a university in a foreign country (e.g. Cambridge University in the United Kingdom) are not part of our study.

C Degrees, country association and employer classification

In the context of this analysis, all of the first university degrees earned by the current and former doctoral students in our sample, as indicated in their respective LinkedIn profiles, were considered bachelor's degrees. 'Diplom' degrees have been included in our analysis as both bachelor's and master's degrees. Degrees earned in foreign countries have been categorized according to the bachelor and master system to the best extent possible.

In some profiles, people indicated several bachelor's and master's degrees. In such cases, we have taken into consideration the first bachelor's degree and the latest master's degree. If a LinkedIn profile shows several doctorates, we have taken into consideration only the most recently completed or the current doctorate. Our analysis did not include any internships, exchange programmes or semesters abroad. We have analyzed the current employment of individuals only if they completed their doctorate already.

The locations of educational stages is not indicated on LinkedIn. Therefore, we assigned a country based on the name of the educational institution indicated in the profiles. About 80 per cent of the institutions could be attributed to a country with the help of the database Research Organization Registry (ROR). Another 14 per cent of the educational institutions were matched using an API of the company OpenCage, which is based on various open sources and tools for georeferencing. All educational institutions that could not be georeferenced using the ROR database or the OpenCage API were manually matched to a country. The automatic connection of countries to educational institutions through the ROR database and the OpenCage API was manually checked on a random basis for 200 cases. This check revealed an accuracy of 99 per cent.

In contrast to the information on educational stages for which LinkedIn does not provide the option to specify a location, LinkedIn users can enter information about their place of work in a text box. For 93 per cent of the LinkedIn profiles on which our analysis is based, a location has been indicated in a text box, which we used for the country attribution. For professions or employers where no such location has been indicated, we linked the information to the respective countries through the OpenCage API or the ROR database (in the case of research institutions). The other professions or employers that could not be georeferenced in this way have been manually matched to a country.

We performed manual online research in order to classify employers into private sector and academic institutions.

Annex 2: Description of the data basis

A Completeness of the data

Since the information in the underlying LinkedIn profiles was not always complete, our findings are based on different sample sizes. In total, we identified the names of 1557 current and former doctoral students and found LinkedIn profiles for 898 of them. Information on the bachelor's degree is available for a total of 727 individuals (see table 2). According to their self-descriptions on LinkedIn, 459 individuals have not yet completed their doctorate. Of the 439 individuals who, according to their profile, completed their doctorates, most of them (396 persons) also indicated their current employer. For 305 of these 396 individuals, information on the bachelor's degree is also available. These 305 individuals form the basis for our analysis on the mobility of former top doctoral students (see figure 5).

Table 2

		Information on the bachelor's degree		
		Information on the bachelor's degree available	Information on the bachelor's degree not available	**Total**
Information on current employment	Doctorate in progress	405	54	459
	Doctorate completed & information on current employer available	305	91	396
	Doctorate completed & information on current employer not available	17	26	43
	Total	727	171	**898**

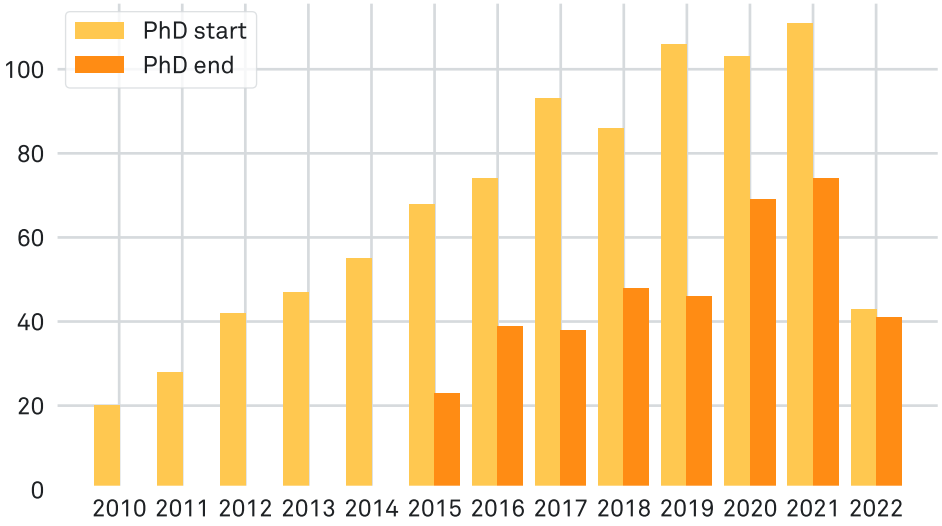
The table gives an overview of the completeness of LinkedIn profiles that form the basis of our analysis. For more than 80 per cent of the total of 898 profiles, information on the bachelor's degree is available. For 95 per cent, there is information on the current stage (doctorate in progress or current profession). For 710 individuals, i.e. 79 per cent of the profiles, both data are available. About half of the former doctoral students in our sample have already completed their doctorate.

B Distribution of PhD start and completion years

For our dataset, we have collected names of doctoral students as they were indicated at the respective time (2022/2018/2015) on the (archived) personal or university websites. These websites do not provide information about the start or termination date of the doctorates. This information was, however, often available on LinkedIn: 98 per cent of the 898 LinkedIn profiles of doctoral students contain information on the start of the doctorates, and 86 per cent on their expected termination. Figure 9 illustrates the distribution of start and termination years: On average, the doctoral students in our sample started their PhDs between 2017 and 2018. On average, the individuals who have already completed their doctorate completed it in 2019.

Figure 9

When did doctoral students in AI included in our dataset start and finish their doctorates?



The histogram shows the number of doctoral students in our sample per year who started their doctorates (light color/left, n=876) and completed them (dark color/right, n=378). The histogram is based on all doctoral students in our sample whose LinkedIn profiles contain information on the period of their doctoral studies.

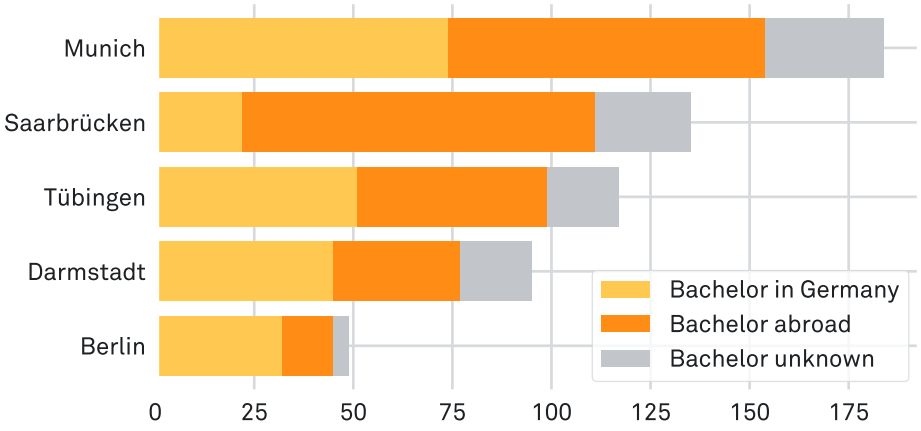
The y-axis shows the absolute number of individuals. Among those who have already completed their doctorates, the doctorate termination date was, on average, 3 years ago (standard deviation +/-2 years). Since our data was collected in mid-2022, the absolute figures for this year are lower.

C Top PhD locations of identified doctoral students

In our sample, most of the top doctoral students in AI are located in Munich, followed by Saarbrücken and Tübingen (figure 10). Within Germany, locations differ with regard to their internationality and the number of doctoral students who study for their doctorate in research groups of top researchers (top doctoral students). At the two universities in the city, Munich trains a particularly large number of top PhD students who are supervised by top researchers. In Munich, the proportion of doctoral students who earned their first university degree in Germany is around 50 per cent. Saarbrücken is a particularly attractive location in Germany for international doctoral students: almost 80 per cent of the top PhD students in Saarbrücken have not earned their bachelor's degree in Germany. Compared to other locations in Germany, Berlin has relatively few doctoral students who have earned their bachelor's degree in a country other than Germany according to our data.

Figure 10

At which locations do top AI talents study for their doctorates in Germany?



The figure shows the five cities with most of the current and former top doctoral students which together represent 66 per cent of our sample. The colors within the bars show the proportion of people at the respective locations with bachelor's degrees earned in Germany, a foreign country or somewhere unknown. The y-axis shows the absolute number of persons. In our sample, most of the top doctoral students in the AI field are located in Munich, followed by Saarbrücken and Tübingen. These cities also have the most top researchers throughout Germany. The proportion of international doctoral students varies according to location: In Saarbrücken, many people do their doctorates who earned their first university degree in a country other than Germany; in Berlin, however, this proportion is significantly lower.

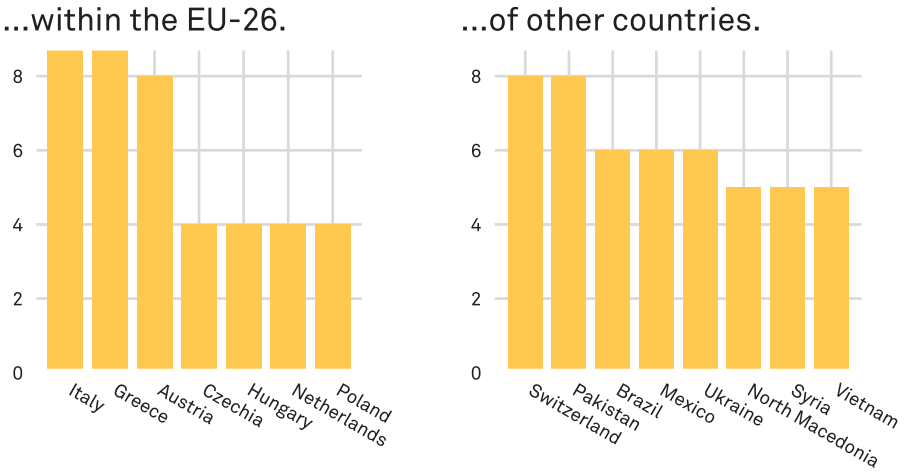
Annex 3: Addendum to figures in the main section

A Countries of first university degree (figure 2)

Figure 2 illustrates the most common countries of origin (countries in which the first university degree was earned), which together represent 87 per cent of the sample. The countries of the European Union are grouped together as EU-26 (excluding Germany). The left diagram of figure 11 shows in which countries of the EU more than 4 individuals have earned their first university degree prior to moving to Germany in order to earn their doctoral degree. The countries of the EU excluding Germany represent a total of 9 per cent of our sample. The right diagram of figure 11 shows the other countries that have not been indicated in figure 2 and in which more than 5 individuals have earned their first university degree prior to moving to Germany in order to earn their doctoral degree.

Figure 11

Breakdown of the first university degree among AI Ph.D. students...



The figures show the countries where the current and former doctoral students included in our sample earned their first university degree and that are either combined in the category 'EU-26' in figure 2 (left) or are not indicated in figure 2 (right). Thus, these figures contain additional information to figure 2. The y-axis shows the absolute number of individuals. Within the category 'EU-26', most of the individuals in our sample earned their first university degree in Italy or Greece.

B Mobility overview table (figure 5)

Table 3 shows the combinations of countries of origin and destination included in our sample. For the purpose of the mobility diagram (figure 5), all individuals with regard to whom information is available on both their bachelor's degree and their employment subsequent to the completion of their doctorate in Germany were considered. As a result, the column and row contained in this table with the headlines 'no information on the bachelor's degree' or 'no information on the current employer' have not been included in the mobility diagram. For that reason, the total numbers in the table differ from those in figure 5.

Table 3

		Destination country								
		Germany	EU-26	Canada	Switzer-land	UK	USA	Other desti-nation countries	No infor-mation on the current employer	Total
Country of origin	China	4	1	0	4	0	2	3	3	17
	Germany	113	8	1	11	6	13	1	11	164
	EU-26	18	4	1	2	2	3	0	1	31
	India	14	2	0	1	2	8	1	1	29
	Iran	10	0	1	0	0	1	0	1	13
	Russia	4	0	0	1	0	1	1	0	7
	Turkey	2	0	0	1	0	1	1	0	5
	USA	3	0	0	0	0	3	0	0	6
	Egypt	2	0	0	0	2	0	1	0	5
	Other countries of origin	29	2	0	3	3	6	2	0	45
	No infor-mation on the bachelor's degree	50	8	4	4	6	14	5	26	117
	Total	249	25	7	27	21	52	15	43	439

Combinations of countries of origin (rows) and destination (columns) among the former doctoral students in our sample. For example, of the total of 29 graduates of a German doctoral programme with an Indian bachelor's degree, 14 work in Germany and 8 in the USA.

Annex 4: Self-selection effects and control estimates

Our analyses presented here are based on publicly available resume data, which have been voluntarily provided on LinkedIn. Like surveys, self-selection effects may occur. This is the case when the availability of the information correlates with the answer to the question. For example, in our analysis there could be a bias if the probability of having a LinkedIn profile (and providing information that is relevant for us in the profile) is related to the country of origin, destination country or employer. We have found LinkedIn profiles for 58 per cent of the total of 1557 top doctoral students identified; 80 per cent of these profiles contain information on the bachelor's degree (cf. table 2).

It is possible that individuals from certain countries may be more likely to have incomplete or lacking LinkedIn profiles compared to those from other countries. To better assess the impact of sampling on our findings, we have used a proxy for a control estimation. For the cohort for which data on the bachelor's degree country was not available, we have estimated the origins of the names and used them as a proxy for the bachelor's degree countries. For that purpose, we used a proprietary API of the French software company Namsor that permits estimations based on names and surnames.

We supplemented the data related to 830 individuals for whom no information on the bachelor's degree was available (due to the absence of LinkedIn profiles or incomplete information in the LinkedIn profiles) with the Namsor estimates and analyzed them together with the 727 complete LinkedIn profiles (n=1557). The findings obtained in this way do not differ significantly from the findings that are based only on the 727 LinkedIn profiles that contain information on the bachelor's degrees. The core finding remains the same: China, India, Iran and Russia are, in addition to Germany, the countries in which most of the bachelor's degrees were earned and the magnitudes also remain the same. Thus, it is unlikely that the findings regarding the first university degrees in our data are heavily biased by the uneven distribution of LinkedIn profiles in our sample.

We do not have such a control tool with regard to the findings on target countries and employers, and as a result, self-selection effects may not be ruled out. It is possible, e.g., that LinkedIn is used to a different degree among employees in the private-sector and in the academic sector.

About Stiftung Neue Verantwortung

Stiftung Neue Verantwortung (SNV) is a non-profit think tank working at the intersection of technology and society. SNV's core method is the collaborative development of policy proposals and analyses. SNV experts do not work on their own, but rather develop and test their ideas together with representatives of politics and public administration, tech companies, civil society and science. The work of our experts is independent from any lobby groups and political parties. We ensure our independence by means of mixed financing with contributions from numerous foundations, public funds and corporate donations.

About the authors

Pegah Maham is lead data scientist at Stiftung Neue Verantwortung, where she has been building the data science team since 2020 and integrated data science into the think tank's work. In particular, she is concerned with the question of how trends in the development of artificial intelligence and its social consequences can be captured using data science methods. She is a co-founder of the Policy Data Science Network, which connects data scientists in public policy, think tanks, and nonprofits. Previously, Pegah Maham worked as a data scientist at a startup in the renewable energy sector. She holds master's degrees in Statistics (Humboldt University of Berlin) and Data Science (London School of Economics).

Dr. Stefan Heumann is a political scientist. Before his appointment to the managing board, he founded and managed the programme 'European Digital Agenda' until March 2016. In this way, he created the basis for the further development of Stiftung Neue Verantwortung (SNV) into a think tank at the intersection of technology and society. Stefan Heumann works, speaks and writes on topics of national and international digital politics. He serves as an honorary board member of the Open Knowledge Foundation Deutschland e. V. and served as an expert member of the Enquete Commission 'Artificial Intelligence' of the German parliament (from September 2018 to October 2020).

Wiebke Denkena is project manager for technology and geopolitics with SNV. Her work is dedicated to the social, (geo)political and economic impacts of artificial intelligence. She holds a bachelor's degree in geodesy and geoinformatics and is an alumni of the interdisciplinary 'Studium Individuale' programme at Leuphana University in Lüneburg. She completed her master's degree in the research-oriented study programme 'Media Studies (Research)' with a specialization in new media and digital cultures at the University of Amsterdam.

Laurenz Hemmen is a data scientist with SNV and is, first of all, in charge of technical aspects of data collection, preparation, analysis and visualization. Before joining SNV, he conducted research on machine learning applications in cancer therapy at the Max Delbrück Center for Molecular Medicine. Laurenz Hemmen studied physics (M.Sc.) at the Georg-August-Universität Göttingen. He focused on machine learning and artificial neural networks, with research stays covering the topics computer vision and natural language processing in India and Taiwan. He is particularly interested in the safety and interpretability of artificial intelligence and also dealt with these issues in his master's thesis.

Anna Semenova works as data scientist with SNV. Their expertise includes data collection and analysis as well as modern practical methods of statistical inference and econometrics. Anna Semenova earned a bachelor's degree in economics from the Higher School of Economics and the New Economic School in Moscow as well as a master's degree in statistics from the Humboldt University of Berlin. Their master's thesis focused on multivariate biomedical data in order to identify patterns for imputing missing values and the translation of a theoretical statistical model into reproducible pipeline code. Before joining SNV, Anna Semenova worked for several companies and gained practical experience in various fields of data science.

Author contact information:

Pegah Maham
pmaham@stiftung-nv.de

Dr. Stefan Heumann
Member of the Managing Board
sheumann@stiftung-nv.de

Wiebke Denkena
wdenkena@stiftung-nv.de

Laurenz Hemmen
lhemmen@stiftung-nv.de

Anna Semenova
asemenova@stiftung-nv.de

Imprint

Stiftung Neue Verantwortung e.V.
Beisheim Center
Berliner Freiheit 2
10785 Berlin

T: +49 (0) 30 81 45 03 78 80
F: +49 (0) 30 81 45 03 78 97

www.stiftung-nv.de
info@stiftung-nv.de

Design:
Make Studio
www.make-studio.net

Layout:
Alina Siebert
Jan Klöthe