



DAAD RESEARCH BRIEF

March 2023

Lessons learnt from a bibliometric study of DAAD scholarship holders

ABSTRACT

In this bibliometric analysis, the authors investigate the feasibility of collecting publications by DAAD-funded researchers and they explore insights into publication patterns that can be gathered from the outputs. While it was possible to extract a considerable set of publications, this process revealed the cruciality of unique author and publication identifiers. Regarding the affiliations of the funded researchers, an increase in connections with Germany can be observed around the funding period, with a large share continuing beyond the funding period. At the same time, German collaborators also establish additional links with sending countries as well. Finally, the DAAD funding can be linked to researchers and their work from a diverse set of countries, and various fields of science, although with a strong orientation towards the biomedical and health sciences and other natural sciences.

<https://doi.org/10.46685/DAADStudien.2023.04>

Introduction

This summary report condenses the most important results of a larger study on the potential of bibliometric analyses for exploring and describing the publication activity of DAAD-funded researchers. The main aim was to conduct several descriptive analyses on the corpus of outputs published by a set of DAAD-funded researchers in order to learn more about their publication activities, in particular in relationship to the period of funding by DAAD.

AUTHORS



Prof Johann Mouton is the director of the DSI-NRF Centre of Excellence in Scientometrics and STI Policy, and Professor at the Centre for Research on Evaluation, Science and Technology (CREST) at Stellenbosch University (South Africa).



Dr Rodrigo Costas is a senior researcher at the Centre for Science and Technology Studies (CWTS) at Leiden University (the Netherlands), and Extraordinary Associate Professor at the Centre for Research on Evaluation, Science and Technology (CREST) at Stellenbosch University (South Africa).



Mr Jonathan Dudek is researcher at the Centre for Science and Technology Studies (CWTS) at Leiden University (the Netherlands).

The study focused on a set of 4,802 individual researchers who received a DAAD scholarship (subsequently referred to as ‘funding’) between the years 2012 and 2018. The candidates came from more than 140 countries and studied – for shorter and longer periods of time – at more than 300 universities and research institutes in Germany. The average age of these candidates was 35.5 years in 2020. Candidates were carrying out research work across more than 200 different scientific fields.

In this summary report, we focus on four specific aspects in the main study:

1. the importance of ORCID and publication identifiers for the proper identification of the funded researchers and their outputs;
2. the development of the publication activity and the affiliation dynamics of the funded researchers;
3. the collaboration networks derived from the publications of the funded researchers; and
4. the topical and geographical diversity of the publication output of the funded researchers.

Methodology

Finding and identifying the publications of the DAAD-funded researchers was a rather challenging process that required multiple steps and data sources. A first step consisted of the algorithmic matching of the names of the funded researchers with author names existing in a database like Web of Science (WoS). This matching was performed with the help of an algorithm that automatically matches authors with sets of publications based on criteria such as names and addresses. In a second step, the output from this automatic matching process was further validated by human coders. A third step consisted of the manual compilation of additional author information like ORCID identifiers, LinkedIn accounts and other websites (such as Google Scholar and Research Gate) that could refer to the identities and the outputs of the funded researchers. A fourth step focused on running queries in other global databases like ORCID and the publication databases Dimensions and Scopus in addition to WoS in order to expand the author and publication information. Finally, a consolidated dataset combining the overall set of publications and the author information

provided by the different databases was created. This rich dataset was used in the subsequent analysis to provide a comprehensive picture of both the output and the main academic dynamics (e.g., collaboration networks and affiliations) of the DAAD-funded researchers.

Central findings

IMPORTANCE OF ORCID AND PUBLICATION IDENTIFIERS

The identification of the academic outputs of the individual researchers was a fundamental step in the study. Like in any bibliometric approach focused on the activities of individual researchers, limitations related to the ambiguity of author names, like homonymy and synonymy, represent formidable obstacles to the correct identification of their full scholarly outputs. In our study, the procedure for identifying and collecting publication data related to the DAAD-funded researchers greatly benefited from the existence of ORCID identifiers. The use of ORCID was of major help not only for unambiguously identifying DAAD-funded researchers, but also for correctly connecting the researchers with their author and publication information across different databases. Relying solely on publication-based approaches (i.e., identifying the individual researchers only based on the bibliographic information of their publications) would have significantly restricted the identification of researcher output.

Table 1 shows that approx. 48 percent of the funded researchers could reliably be associated with an ORCID profile. Interestingly, over 66 percent of all the DAAD-funded researchers could also reliably be associated with a publication indexed in any of the publication databases considered. However, out of the 3,677 researchers with with a publication or a researcher ID, there was a substantial number of researchers (1,392 or 38%) who produced publications but did not present an identifiable ORCID profile. Moreover, during the manual verification of the data recorded for this project, the manual coders encountered challenges related to the actual use of ORCID profiles by researchers. Sometimes, ORCID profiles came with very low degrees of completeness (e.g., lacking biographical or affiliation data),

making the proper identification of the individual researcher to whom the ORCID belongs unfeasible.

These experiences suggest that the adoption of ORCID by DAAD-funded researchers could be more broadly encouraged as the most reliable approach to identify and make visible their academic activities and outputs. DAAD could also implement trainings for funded researchers on best practices on how to manage and maintain ORCID profiles in order to extract the most value from them.

TABLE 1: COVERAGE OF DAAD-FUNDED RESEARCHERS

Type of identifier	Counts	Percentage of total
Total number of DAAD funded researchers	4,802	100%
Researchers with publication or researcher ID	3,677	76.6%
Researchers with publication ID	3,183	66.3%
Researchers with assigned DOI	3,129	65.2%
Researchers with an ORCID identifier	2,285	47.6%

3

In line with the importance of ORCID, the methodological approach for this study also benefited from the existence of publication identifiers, particularly from Digital Object Identifiers (DOIs). DOIs proved to be critical for tracking publications across different databases. In the future, DAAD should encourage the use of this type of identifiers. Together with ORCID, they allow the outputs of funded researchers to be monitored more comprehensively.

PUBLICATION ACTIVITY & AFFILIATION CHANGES OF DAAD-FUNDED RESEARCHERS

Our study shows that there is a positive link between funding by DAAD and the productivity of the funded researchers, particularly in the years towards the end of the funding period. The highest

productivity is observed in the years immediately after the end of the funding, suggesting some sort of latency of the activities and publications of the researchers after the funding has finished. This simply means that most students only start publishing after completion of their studies and not whilst they are still engaged in their doctoral studies. Another explanation could be that funded individuals may continue their research activities beyond the DAAD-funded period. These aspects would clearly benefit from future research, particularly using additional qualitative approaches (e.g., interviews or surveys) since bibliometric approaches may be limited in capturing the future professional developments of the funded researchers once their funding is over.

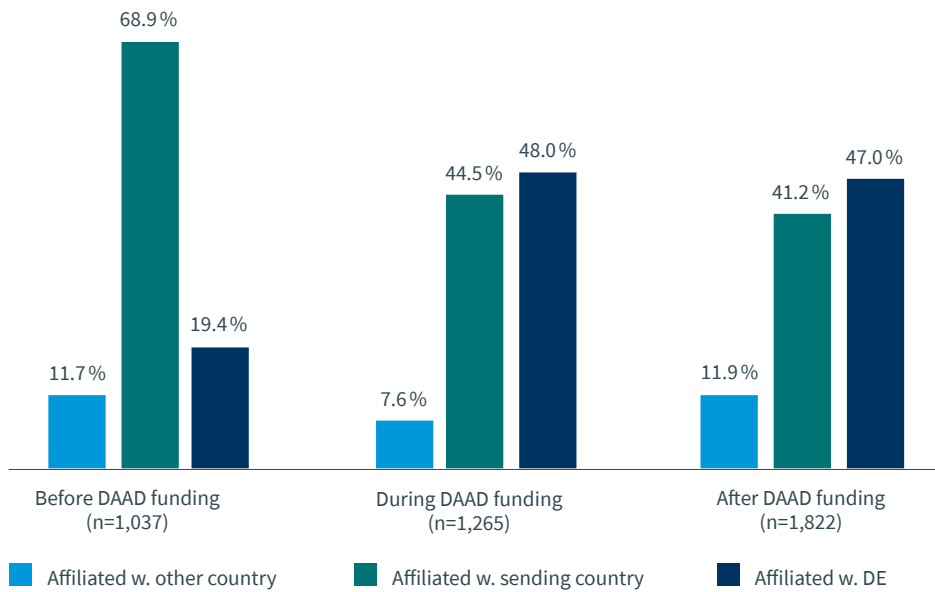
Beyond productivity, insights regarding changes in the affiliations of the funded researchers are also possible. Researchers indicate their affiliations in their publications, which makes it possible to analyse changes in their affiliations over time. This information tells us which countries the researchers were working in before they received DAAD funding, who they were affiliated with during their funding, and where they continued their research afterwards. We focused, in particular, on an analysis of the researchers' countries of affiliation and classified these into three groups:

- countries targeted by DAAD ('sending countries'),
- Germany (as the main country in which DAAD-funded researchers are expected to stay during the funding period),
- other countries (i.e., countries other than sending countries or Germany).

Figure 1 shows the percentages of researchers per period and country type, based on affiliation information extracted from WoS.¹

¹ Researchers can have multiple affiliations (in a given publication, but also in a given period). When researchers have multiple affiliations within one of the three periods, affiliations are counted as fractions. The sums of the percentages per period are rounded up in Figure 1 and can therefore be slightly above 100 percent.

FIGURE 1. SHARES OF DAAD-FUNDED RESEARCHERS BY AFFILIATION, ACCORDING TO PUBLICATION ACTIVITY IN WOS



Note: N = 2,114

4

The share of researchers affiliated with Germany increases considerably during the funding period.² Interestingly, a large share of the funded researchers continues to have an affiliation in Germany after the DAAD funding was ended, which may suggest that these researchers have established further research relationships in the country (Germany) beyond the funding period.

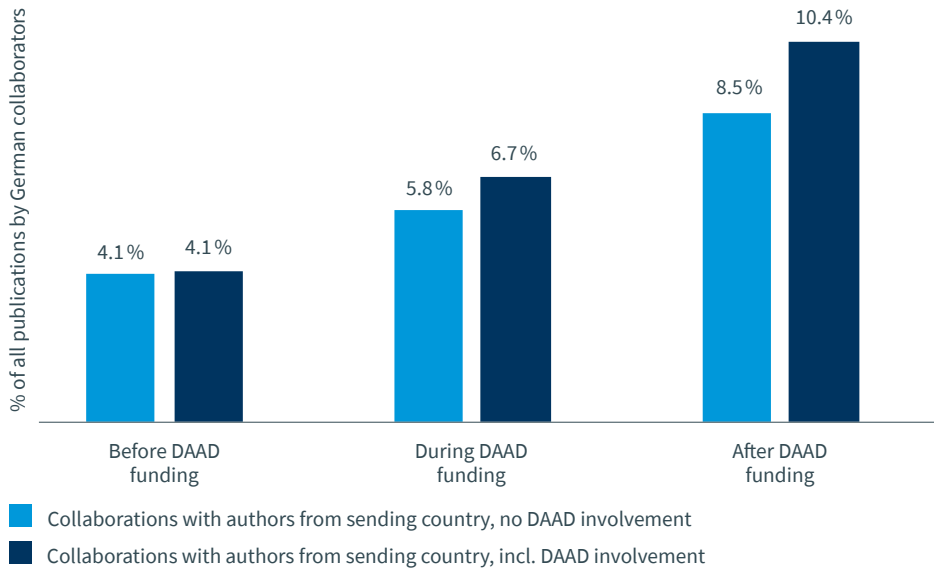
COLLABORATION NETWORKS BY GERMAN CO-AUTHORS

The implications of funding through DAAD might go beyond just the funded researchers. More specifically, it might increase the collaboration networks between researchers in Germany and researchers in the sending countries, sometimes even without

the participation of the funded scholars. In order to test this point, we looked at the publications of the German co-authors of the DAAD-funded researchers and compared their collaboration ties with the sending countries before and after the arrival of the DAAD-funded researchers. We focused on the comparison of the publications that were co-authored by the German collaborators and a DAAD-funded researcher, and publications co-authored by the German collaborators, but without the participation of DAAD-funded researchers. For both these sets we calculated the share of publications involving a co-author from a sending country. Again, we also applied the three different time periods (before, during, and after the funding). The results are shown in Figure 2.

2 In the case of researchers who came to Germany for more than one year, the share affiliated with Germany also continues to increase after the funding period.

FIGURE 2. COLLABORATIONS WITH AUTHORS FROM SENDING COUNTRIES AMONG PUBLICATIONS OF GERMAN COLLABORATORS OF DAAD-FUNDED RESEARCHERS.



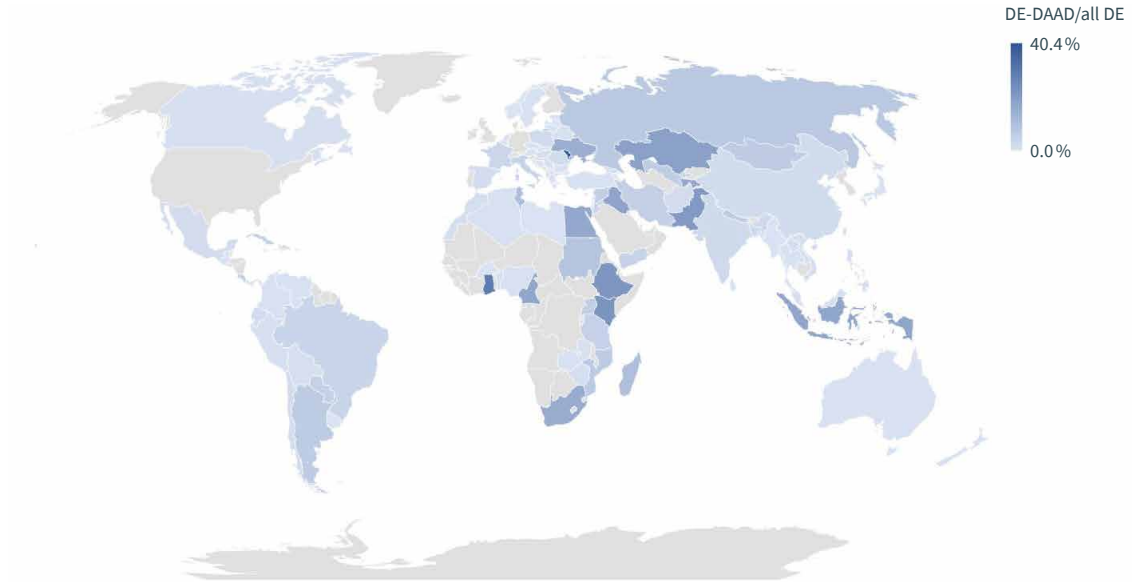
Source: Based on publications in WoS

5

The share of collaborations of German researchers with researchers from the sending countries increases during the funding period and continues to do so after the funding period. This applies both to the situation where we included publications co-authored by DAAD-funded scholars (dark blue bar), but also where the DAAD-funded scholars are not included (light blue bar). This suggests the potentially beneficial role of DAAD funding for both German researchers and researchers from the sending countries with a mutual increase in collaborations. Whereas the DAAD-funded scholars intensify links with Germany in particular, the German collaborators also increase collaborations with other researchers from the sending countries.

To assess the role of the DAAD funding for collaborations, we can also look at all the publications that were co-authored by researchers from Germany and a sending country. Among those publications, we can calculate the share of publications that involved a researcher from Germany who has collaborated at some point with DAAD-funded researchers. Here, some countries stand out, e.g. Moldova (40.4%), Ghana (29.0%), Kenya (22.3%), Ethiopia (21.8%) or Kazakhstan (18.8%). The role of the German collaborators of DAAD-funded researchers among all collaborations with Germany is particularly significant in those countries (see also Figure 3).

FIGURE 3. WORLDWIDE SHARES OF COLLABORATIONS BETWEEN SENDING COUNTRIES AND GERMAN DAAD-COLLABORATORS, AND ALL COLLABORATIONS BETWEEN SENDING COUNTRIES AND GERMAN RESEARCHERS.



Source: Based on publications in WoS

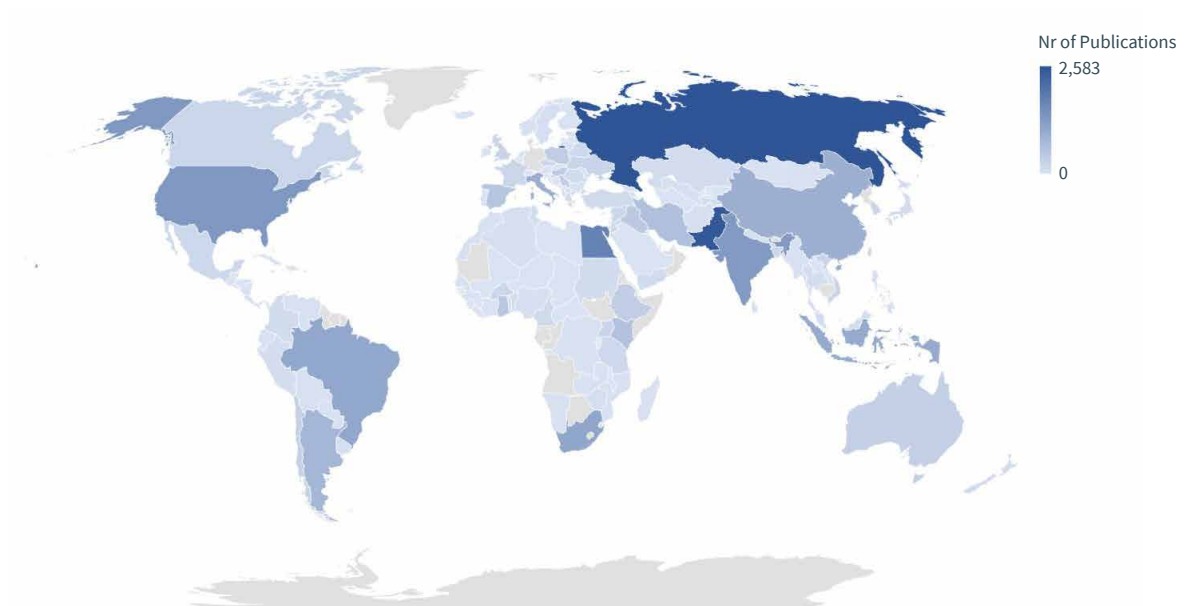
6

GEOGRAPHICAL AND TOPICAL DIVERSITY OF DAAD-FUNDED RESEARCHERS

The DAAD funding focuses on a range of countries. This can be characterised in different ways, but one is to connect publications (co-)authored by the DAAD-funded researchers to their sending countries. When considering the publication output in

absolute numbers, several countries stand out, e.g., Russia (2,583 publications), Pakistan (2,506 publications), Egypt (1,748 publications), the United States (1,355 publications), India (1,329 publications), South Africa (1,095 publications), Brazil (1,072 publications) or Indonesia (1,010 publications). Figure 4 visualizes the global distribution.

FIGURE 4. PUBLICATION OUTPUT (BY DOIS) PER SENDING COUNTRY.

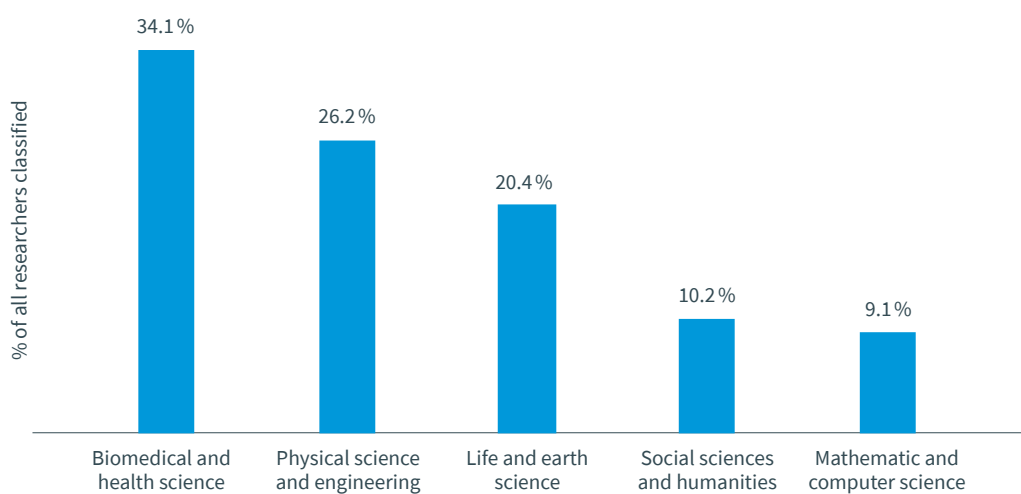


Source: Based on DOIs across all databases

DAAD classifies the funded researchers according to fields of research. With the collected publication information, however, the researchers can also be linked to classifications that are based on publication output. One such classification is used in the Leiden Ranking produced by CWTS.³ Focusing on the publications found in the Dimensions database,

we could link all publishing researchers to a field.⁴ Figure 5 presents the percentages of researchers for each of the five fields of science included. The biomedical and health sciences take the largest share of researchers with 34 percent, and mathematics and computer science the lowest share, with nine percent.

FIGURE 5. PERCENTAGES OF RESEARCHERS PER FIELD OF SCIENCE, LEIDEN RANKING CLASSIFICATION SYSTEM



Source: Based on publications found in the Dimensions database

Main lessons learnt and future research

The work reported presents various lessons learnt. A general lesson learnt is that bibliometric analysis can provide funders like DAAD with unique information about the characteristics, dynamics and developments of the individual researchers they have funded. The intelligence extracted from such bibliometric analyses can go beyond the mere counting of publications (or citations) and reveal other relevant aspects such as the diversity of topics, geographical connections developed by the funded researchers, or the collaboration networks that have arisen from the funding. All these are relevant aspects that can inform the current and future funding policies of a funder like DAAD.

Against this background, it is important to highlight that the process of identifying the bibliographic data of the DAAD-funded researchers and, consequently, their academic output, was rather complex and required the use of multiple data sources. This is certainly an aspect that should be improved in the future, particularly by adopting workflows that simplify the reporting and identification of outcomes by DAAD-funded researchers. Below, we list some of the key lessons from the bibliometric study conducted.

³ <https://www.leidenranking.com/information/fields>

⁴ Researchers were linked to the fields on the basis of publication counts: each researcher was assigned to the field of activity which recorded the highest number of publications. Because a research can in principle be assigned to different fields depending on the journals in which they publish, it is possible that a specific researcher could be assigned to two different fields with a similar share (this is referred to as a 'tie'). Whenever ties in the top fields of researchers by the number of publications existed, tiebreakers were applied in the following order: 1) the field with the highest number of citations, 2) the field with the largest number of co-authors, 3) the field with publications in the most recent years. If ties still remained in the profile of a researcher, one of them was chosen randomly.

BROADER ADOPTION AND USE OF UNIQUE IDENTIFIERS

The adoption and use of ORCID by all researchers funded by DAAD is the way forward in order to systematize the identification and reliable documentation of their research activities and outcomes. This is the most effective way to properly identify them and make sure that their production and affiliations are not missed. It is also important to highlight that the creation of an ORCID profile alone is not sufficient to properly identify the outputs of funded researchers. ORCID records need to be maintained and populated, e.g. with affiliation information, research outputs, and academic activities.

At the same time, researchers should make sure that their outputs can be identified and traced across databases in reliable ways. This can be achieved by ensuring that identifiers like DOIs (preferably) are provided for documents. The lack of identifiers makes the accountability of scientific results more complicated and prone to error (e.g. double counting or underreporting).

AFFILIATIONS

As the results have shown, funded researchers tend to maintain affiliation links with German organizations even after the conclusion of DAAD funding. This suggests that funding contributes to lasting, sustainable research networks. Future research should explore how DAAD funding may help

foreign researchers establish longstanding research affiliations with Germany.

COLLABORATIONS

Results provide evidence that DAAD funding may be beneficial not only to the researchers, but also to their collaborators in Germany as well as other researchers in their country of origin. The results support the idea that expanded collaboration networks are developed before, during and after funding. We therefore strongly recommend that further research should be carried out regarding the network spillover effects of funding like those from DAAD.

GEOGRAPHICAL AND TOPICAL DIVERSITY

The analysis shows that publication output by DAAD-funded researchers can be linked to a broad set of countries, especially Russia, Pakistan, Egypt, and the United States. Other notable examples are India, South Africa, and Brazil.

Regarding topical diversity, DAAD-funded researchers could be classified according to five broad fields of science based on their scientific output. More than one third of the funded publishing researchers can be related to the biomedical and health sciences, approx. one quarter to the physical sciences and engineering, and one fifth to the life and earth sciences. Only about nine percent of the funded scholars can be related to the social sciences and humanities, and mathematics and computer sciences.

IMPRINT

Editor
Deutscher Akademischer Austauschdienst e.V. (DAAD)
Kennedyallee 50, D-53175 Bonn,
Tel.: +49 228 882-0, Fax: +49 228 882-444
E-Mail: webmaster@daad.de, Internet: <https://www.daad.de>
Authorised Representative of the Executive Committee:
Prof Dr Joybrato Mukherjee, District Court of Bonn,
Register of associations, number VR 2107,
Sales tax number: DE122276332
Person responsible according to § 18 Abs. 2 MStV:
Dr Kai Sicks, Kennedyallee 50, 53175 Bonn

Contact
Division Strategic Planning – S1
Dr Jan Kercher, kercher@daad.de

Layout and typesetting:
Atelier Hauer+Dörfler GmbH, Berlin

Proofreading
cleevesmedia e.K., Meckenheim

March 2023
Published as a digital publication on the internet
This work is licensed under the Creative Commons licence
Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0)
<https://creativecommons.org/licenses/by-sa/3.0/>

