

Phases of Development

BiCDaS Activity Report – 2018/19



Image authorship for cover image see Impress.

Images on the cover:

Array of identical memory units (second from bottom): © Björn Fischer
All other images: © Bielefeld University

Impress

Bielefeld University
Bielefeld Center for Data Science (BiCDaS)
Universitätsstraße 25
33615 Bielefeld
Germany

Phone: 0049 521 106 65799

bicdas@uni-bielefeld.de

www.uni-bielefeld.de/datascience

Content and Design

Bielefeld Center for Data Science (BiCDaS)

Nils Hachmeister und Reinhold Decker

bicdas@uni-bielefeld.de

Content

| | |
|--|----|
| Part A BiCDaS Portrait | 1 |
| A - 1 BiCDaS Goals and Context | 1 |
| A - 1.1 Four Pillars of Activity | 1 |
| A - 1.2 Horizontal Topics/Guiding Principles | 1 |
| A - 1.3 Our Academic Context – Bielefeld University | 2 |
| A - 1.4 Academic Partners | 2 |
| A - 2 Structure | 4 |
| A - 2.1 Round Table Data Science (BiCDaS General Assembly) | 4 |
| A - 2.2 Staff Members..... | 4 |
| A - 2.3 Scientific Advisory Board..... | 5 |
| Part B BiCDaS up to 06-2019..... | 5 |
| B - 1 Activities | 5 |
| B - 1.1 Data Science Teaching..... | 5 |
| B - 1.2 Interdisciplinary Data Literacy Module | 6 |
| B - 1.3 Data Literacy Education Network..... | 6 |
| B - 1.4 Call for Data Literacy Teaching..... | 6 |
| B - 1.5 Lecture Series Data Science | 7 |
| B - 1.6 Data Science Day | 7 |
| B - 1.7 Co-Organisation of Scientific Events | 7 |
| B - 1.8 Support with Funding Proposals | 8 |
| B - 1.9 Research Data Management..... | 8 |
| B - 1.10 International Networking in Science - EuADS | 8 |
| B - 1.11 Exchange with Companies | 9 |
| B - 1.12 Exchange with Faculty of Medicine..... | 9 |
| B - 1.13 Support of GPU-HPC cluster..... | 9 |
| B - 1.14 Coordinating Activities in Context with NFDI..... | 9 |
| B - 2 Communication | 10 |
| B - 2.1 Internal Communication | 10 |
| B - 2.2 External Communication..... | 10 |
| B - 3 Future Plans/Strategies | 11 |
| B - 3.1 Interdisciplinary Dialog (DataLabs) | 11 |
| B - 3.2 Stimulate Funding Proposals..... | 11 |
| B - 3.3 Using Momentum | 11 |
| Part C – Conclusion..... | 11 |

Executive Summary

The **Bielefeld Center for Data Science** is a highly interdisciplinary network at **Bielefeld University**, bringing together researchers from all faculties and several central institutions. It was initiated by the Vice-Rector for Information Management and is supported by the rectorate of Bielefeld University. The centre is developed by an executive director who reports to the general assembly of BiCDaS members.

The approach to data science pursued at BiCDaS puts forward two aspects of data science:

Interdisciplinarity: BiCDaS aims to initiate and support scientific dialog and exchange, structured along the lines of common data rather than discipline boundaries.

Holistic data science: BiCDaS is concerned with (research) data along the entire data life cycle, starting with a (research) question/problem, through data collection, preparation, analysis and exploitation to data storage, sharing and linking.

At Bielefeld University, BiCDaS is embedded in an environment characterised by interdisciplinarity and innovation, featuring many, diverse academic partner institutions such as the Cognitive Interaction Technology Cluster of Excellence (CITEC), the Centre for Biotechnology (CeBiTec), the Centre for Statistics (ZeSt), the Research Centre on Mathematical Modelling (RCM²) and others.

BiCDaS has identified four focus areas for its mission: (i) supporting data science research, (ii) advertising the academic implementation of data science, (iii) supporting researchers in employing data science and (iv) data science teaching. Additionally, two horizontal topics have been identified in stimulation of data-centred, scientific dialogs and data ethics.

The activities of BiCDaS from June 2017 to June 2019 (the report period) encompass support in funding proposals, support in the organisation of events, developing scientific and economic partnerships, support in the development of a competence centre on research data management, organisation of a data science day, organisation of a lecture series on data science, establishing a data science news ticker, initiation of an ongoing exchange with the Faculty of Medicine, developing an agile, data-centred cooperation format (DataLabs), accompanying the creation of a GPU-HPC cluster at Bielefeld University, coordination of activities at Bielefeld University in context with the NFDI research data infrastructure and an internal call for innovative data literacy teaching concepts. Furthermore, in 2018, a second position in the BiCDaS staff (aside the executive director) was created for the coordination of a data literacy project that was started in early 2019. Finally, at an international level, BiCDaS is involved in developing the European Association for Data Science (EuADS).

Part A BiCDaS Portrait

In this part of the report, we discuss the major factors defining BiCDaS, namely, (i) our approach to data science, (ii) our vision for BiCDaS, (iii) the context BiCDaS resides in and (iv) BiCDaS's structure.

A - 1 BiCDaS Goals and Context

A - 1.1 Four Pillars of Activity

When BiCDaS was founded, four pillars that are guiding and structuring our activities have been defined.

Supporting data science research

We believe that data science is a field of knowledge in its own right and should be ploughed as such and, hence, we support corresponding research activities.

Advertising data science/data awareness

The methods and technologies developed in data science are beneficial for research in general. Therefore, BiCDaS aims to make the idea of data science known and to advertise its elements and philosophy among academics.

Supporting data science practice

BiCDaS supports academics in incorporating data science in their daily work. This is achieved, for instance, by providing data literacy training or by organizing workshops. Furthermore, we support academics regarding data-science-related questions.

Fostering data science teaching

The demand for data scientists in science and industry is enormous and rapidly growing. To meet their social obligations, universities need to offer courses and grant degrees on data science and enshrine data literacy as a general objective in academic education.

A - 1.2 Horizontal Topics/Guiding Principles

Besides these four pillars of activity, two horizontal values/topics guide BiCDaS's work.

Data-Centred, Scientific Dialog

We firmly believe that one of the main benefits emerging from the thorough implementation of data science within an academic context is interdisciplinarity. When aspects of data, such as data type(s), analysis, and curation, are the primary concerns, traditional discipline boundaries are consequentially upstaged. Interdisciplinarity easily gives birth to a mutually fruitful exchange of methods, best practices and experiences. Researchers experience the 'best of all worlds' for their data work in research and teaching. The format of the exchange, be it multi- or omnilateral, is of secondary concern and should be chosen appropriately by the researchers themselves.

Data Ethics

There is an active societal discussion on how data should and can be used, on the ethical implications of data usage and data science research and on how the upcoming technologies in the field will transform societies. At BiCDaS, we welcome these discussions as a necessary part of the development of new technologies and a new field of science. As an academic institution, we have the potential and capacity to go ahead in this discussion.

Position Paper

As a basis for its activities, BiCDaS developed a position statement in 2018 in which the members of BiCDaS developed their common approach to data science. While not wanting to add to the abundance of data science definitions, we still felt the need to define our perspective on a wide and multi-faceted topic. In the second part of this position paper, the structure and context of BiCDaS is laid out. The Part A 'BiCDaS Portrait' of this report largely covers the same content as that of the position paper, but in a more detailed form.

A - 1.3 Our Academic Context – Bielefeld University

Bielefeld University is a medium-sized university with 24,875 students, 2,107 academic staff and 1,380 technical and administrative staff¹. It is currently involved in four SFBs (211, 212, 1283, 1288), houses 15 graduate schools and offers 52 courses on bachelor- and 64 course on master-level as well as 12 PhD courses. It has a very spatially compact campus with all institutions and faculties assembled in one place.

Bielefeld University has promoted interdisciplinarity as one of its high-priority goals earlier than most German universities and, today, the benefits of interdisciplinary research are generally considered self-evident by our researchers. This spirit is a major advantage for achieving BiCDaS's goals.

Furthermore, the increasing significance of research data management was recognized by Bielefeld University at a very early stage. Hence, it was among the first German universities to issue a policy on research data management and has continually pursued this topic ever since. Thus, today Bielefeld University is in a lead position in the field of research data management.

A particularly appealing task for BiCDaS emerged in 2017 when the state government of North Rhine-Westphalia decided to establish a faculty of medicine at Bielefeld University and, subsequently, data science in medicine has been identified as a central topic in education and research.

A - 1.4 Academic Partners

BiCDaS is highly interconnected across Bielefeld University. Here, we will briefly introduce some of our academic partners and institutions in alphabetical order.

A - 1.4.1 Bielefeld IT Service Centre

The Bielefeld IT Service Centre (BITS) provides IT services throughout the university. It provides many services crucial to data science, such as long-term data storage and various server solutions and also hosts HPC resources. It is, therefore, a crucial partner in our mission.

A - 1.4.2 CeBiTec

The Centre for Biotechnology (CeBiTec) is a research hub in life science, biotechnology and bioinformatics. In the last decade, this area has seen a rapid decrease in the cost of its main data measurement technology (gene sequencing) with a corresponding explosion in the amount of data. With this data and the promises that it offered, the field was bound to be successful in data science and indeed, today, the life sciences are in many aspects exemplary in their research data management, computing infrastructure and research software.

At CeBiTec, a research cloud providing cloud-based software applications for various analytic tasks, powered by HPC resources and extensive cloud storage, has been developed. The research cloud services at CeBiTec are included in the German Network for Bioinformatics (de.NBI). The de.NBI research cloud is being administered from CeBiTec and the current de.NBI chair is a BiCDaS member.

A - 1.4.3 Centre for Statistics

The Centre for Statistics (Zentrum für Statistik – ZeSt) at Bielefeld University is an interdisciplinary network of researchers with a statistical focus on their respective research fields that span business administration, psychology and sport sciences, educational sciences, mathematics, health sciences and sociology. Members of the centre have designed a master programme on statistical sciences as well as a master programme on data science (see Section 'B - 1.1 Data Science Teaching'). ZeSt also hosts the Statistical Consulting Centre (Statistisches Beratungszentrum – StBeCe), which offers advice on matters of statistics all around the data life cycle.

¹ According to the Annual Statistical Report of Bielefeld University – year 2018.

A - 1.4.4 Centre for Teaching and Learning

The Centre for Teaching and Learning (Zentrum für Lehren und Lernen – ZLL) offers a comprehensive portfolio of services for lecturers and students. It is an important partner for BiCDaS in data science and data literacy teaching. ZLL and BiCDaS also cooperate when companies approach BiCDaS with requests for the advertisement of positions, traineeships or joint theses, which happens regularly.

ZLL also features the programme ‚richtig einsteigen.‘ (roughly translated as ‘getting a good start’), helping freshmen students mitigate their most common hurdles, which have been identified in acquiring academic-level literal and mathematical competencies. Together with BiCDaS, ZLL developed an additional column of this programme aiming to introduce data competencies into the curricula of the different courses offered by Bielefeld University (see Sections ‘B - 1.1 Data Science Teaching’ and ‘B - 1.8.1 Data Literacy Education’).

A - 1.4.5 CITEC

The Cognitive Interaction Technology Cluster of Excellence (CITEC) is an interdisciplinary research centre on the intersection between computer science, machine learning and robotics on the one hand and psychology, sport science and sociology on the other hand. Hence, it encompasses many data-intensive applications, and many active BiCDaS members are part of CITEC as well.

A - 1.4.6 Competence Centre for Research Data Management

The Competence Centre for Research Data Management, a joint venture of the library and BITS, became operational in late 2018. BiCDaS was involved in the centre’s conception (see ‘B - 1.9 Research Data Management’) and both are now partners in the endeavour to improve data work all around the data life cycle. The centre offers guidance and support on research data management and simultaneously pursues the goal to increase awareness of the importance of research data management throughout the university.

It experiences a huge demand from researchers as many public scientific funding agencies have now incorporated requirements regarding research data management into their terms. Particularly, the services provided in creating data management plans (DMPs) are in high demand.

A - 1.4.7 RCM²

The Research Centre for Mathematical Modelling (RCM²) focusses on the application of mathematical methods within the sciences. Currently, it focusses on mathematical biosciences, mathematical physics, and applied dynamical systems. In the area of data science, RCM² plans to contribute model-based approaches, and thus to contribute to ‘explainable’ data science. Such approaches currently play a decisive role in the evaluation of ‘big data’ in social networks as well as neural, genetic, and metabolic systems, to name just a few.

A - 1.4.8 SFB1288

The DFG-funded interdisciplinary research project ‘Practices of comparing. Ordering and changing the world’ investigates the role of the societal practice of comparing from a variety of scientific perspectives: sociology, anthropology, history, literature studies, political sciences, law and others.

The subproject ‘Data Infrastructure and Digital Humanities’ has a particular affinity to data science. It brings together experts from the fields of information infrastructure, research data management and digital humanities. The subproject enriches the methodology of the SFB by computational methods from the digital humanities while simultaneously ensuring the sustainability of the SFB’s scientific output by the long-term availability and traceability of research data and findings.

A - 1.4.9 University Library

The University Library in Bielefeld nowadays is a 21st century general provider of information services. Aside from the standard library services, it offers a broad variety of digital services. As such, the Publications at Bielefeld University (PUB) system is hosted and developed at the library, which is a catalogue containing classical paper publications as well as research data publications across domains. The PUB system is a central element of the open access strategy of Bielefeld University.² Furthermore, the library's IT team hosts different open source software programmes such as a Gitlab or a Share Latex instance.

A - 2 Structure

BiCDaS has about 50 members who come from all faculties and different central institutions. Most of them are researchers; however, non-scientific staff concerned with data science in their daily work, such as those from the university library and BITS, are also members.

A - 2.1 Round Table Data Science (BiCDaS General Assembly)

The round table data science is where BiCDaS originated from and nowadays it is its steering committee. The round table was initiated in May 2016 by the Vice-Rector for Information Management of Bielefeld University. Since then, it has met a total of ten times. The original purpose of the round table was to bring together researchers and employees of Bielefeld University with an interest in data science and to discuss how this topic is best ploughed at Bielefeld University. The answer given by the round table was the formation of the Bielefeld Center for Data Science.

After an executive director was appointed, the round table has played the role of a steering committee of BiCDaS. It meets two to three times a year upon the invitation of the executive director, who informs the round table about recent developments and presents plans for BiCDaS's future course.

A - 2.2 Staff Members

Until February 2019, the only staff member of BiCDaS has been the executive director. Then, a coordinator 'data literacy education' joined BiCDaS as a second staff member. BiCDaS does not incorporate any formal/legal entity up to now and, thus, its staff members are employed at the Department for Information Management and University Development.

A - 2.2.1 Executive Director

The position of executive director was created as a full-time position by the rectorate of Bielefeld University after the round table data science decided to establish BiCDaS. The task of the executive director is to develop BiCDaS. For that mission, the position has been provided with a great deal of autonomy. The executive director reports to the round table and coordinates the activities with the Vice-Rector for Information Management.

A - 2.2.2 Coordinator Data Literacy Education

In the middle of 2018, BiCDaS and ZLL partnered to develop a concept for a 'Data Literacy Education' project, aiming to equip students with data competencies that will enable them to be maximally effective in future jobs and to participate in societal discourses, which is presumed to become increasingly data-driven (see Section 'B - 1.1 Data Science Teaching'). Since mid-February 2019, the coordinator 'data literacy education' has advanced this project as a BiCDaS staff member and in close cooperation with ZLL and the programme 'richtig einsteigen.' (see Section 'A - 1.4.4 Centre for Teaching and Learning').

² See <http://oa.uni-bielefeld.de/index.html> (in German)

A - 2.3 Scientific Advisory Board

BiCDaS as a horizontal institution stretching across all faculties and disciplines and involving different central institutions faces unique challenges in the university environment. At the same time, this same horizontal structure along with the very popular topic of data science offers unique possibilities that will benefit BiCDaS members and research at Bielefeld University in general.

To fully maximize these opportunities while simultaneously facing the challenges in an adequate manner, it is vital to solicit the advice of external, highly experienced, academic professionals. Given the wide scope of data science as a topic and the multi-faceted BiCDaS members, we aim for the advisory board to mirror these facets and BiCDaS's general philosophy (see Sections 'A - 1.1 Four Pillars of Activity', 'A - 1.2 Horizontal Topics/Guiding Principles').

Part B BiCDaS up to 06-2019

Here, we will briefly report on the activities of BiCDaS during the report period from June 2017 to June 2019, as well as some future plans and strategic considerations.

B - 1 Activities

The years 2017 and 2018 were a time of beginning for BiCDaS. Much effort was spent on networking and (university) public relations. However, we also already contributed to different funding proposals and scientific events. These activities have been and will continue to be intensified in 2019.

B - 1.1 Data Science Teaching

Teaching in the field of data science is one of the four core activities of BiCDaS (see Section 'A - 1.1 Four Pillars of Activity'). In the winter term of 2018/19, Bielefeld University launched a master programme on data science. When BiCDaS was founded in 2017, the programme had already been conceptualized, but since then BiCDaS has supported the master programme.

The master programme is insofar exceptional, as it puts a strong emphasis on balancing statistics and machine learning as the two main methodical sources of the field, with substantial contributions from mathematics on the topic of modelling (see Section 'A - 1.4.7 RCM²'). Consequentially, the programme is supported by the Faculty of Business Administration and Economics, which houses huge competencies in the field of statistics, and the Technical Faculty, complementing this with its equally extensive competencies in algorithms and machine learning, as equal partners.

Since data science teaching was already well underway, BiCDaS could focus on the aspect of data literacy teaching. In June 2018, BiCDaS and ZLL joined forces to participate in a call from the Stifterverband für die Deutschen Wissenschaft³ (roughly translated as Donors' association for the promotion of humanities and sciences in Germany) for innovative and interdisciplinary concepts aiming to incorporate data competencies as a goal for teaching and learning in curricula throughout universities. A multi-faceted team came together to develop the concepts for this proposal: lecturers from different career stages, members of the rectorate, delegates of ZLL and BiCDaS, student representatives and stakeholders from the local business community and society.

This team developed a highly innovative concept. Faculties would be supported in evaluating and advancing their status quo in the teaching of data competencies in an iterative process handling three study programmes per year. Complementarily, an interdisciplinary dialog among lecturers involved in data literacy education would be initiated and supported. Finally, measures were described to increase data awareness throughout the university.

³ See <https://www.stifterverband.org/>

The team of ‘richtig einsteigen.’ already has huge experience in incorporating literal (LitKom) and mathematical (MatKom) competencies in curricula across study programmes (see Section ‘A - 1.4.4 Centre for Teaching and Learning’). To maximally profit from this experience and existing infrastructure, data literacy would constitute an additional pillar (DatKom) within ‘richtig einsteigen.’

Our proposal was ultimately not successful in the highly competitive call, but all stakeholders felt that implementing this concept nonetheless would be to the great benefit of students at Bielefeld University. Therefore, we are most grateful that the project could be funded internally to its full, planned extend.

B - 1.2 Interdisciplinary Data Literacy Module

The data literacy education project ‘DatKom’ aims at integrating data competencies in the curricula of all disciplines at Bielefeld University. Besides this long-term process, BiCDaS together with lecturers from several faculties initiated the development of a sustained data literacy module open to all undergraduate students starting in October 2019.

One part of the module will be the interdisciplinary lecture entitled ‘Big Data is watching you! How to deal with data in the modern world.’ Students will be given an overview about the role of data in science, economy and society, about data analysis and the transformation of data into knowledge. Participants of the lecture can individually choose from a variety of connecting courses that are offered by different faculties to expand on their knowledge, acquire additional data competencies and to complete the data literacy module.



Figure 1 - Promotion for the data literacy lecture ‘Big data is watching you!’

BiCDaS initiated and coordinated the genesis of the data literacy module by bringing together lecturers and the administration and by organizing promotion and evaluation of the lecture.

B - 1.3 Data Literacy Education Network

Data literacy education is being recognized as an increasingly important subject in academic teaching, not only at Bielefeld University but throughout higher education in Germany. Various universities are currently developing individual concepts to provide students of all disciplines with data competencies. To bring together data literacy education protagonists of different universities and to foster the mutual exchange of ideas and concepts, the DATEV Stiftung Zukunft and the Stifterverband für die Deutsche Wissenschaft have founded a national data literacy education network. Led by BiCDaS, Bielefeld University successfully applied for membership in this network. From October 2019, BiCDaS together with ZLL will represent Bielefeld University in the national data literacy education network for a period of two years. The regular exchange and peer consultations among the 14 networking institutions are valuable opportunities for the advancement of data literacy teaching at Bielefeld University.

B - 1.4 Call for Data Literacy Teaching

In the context of the programme ‘richtig einsteigen.’ (see Section ‘A - 1.4.4 Centre for Teaching and Learning’), a yearly internal call for proposals for innovative teaching concepts is issued at Bielefeld

University, the so-called ‘Qualitätsfonds für die Lehre’ (roughly translated as ‘quality fund for teaching’). In 2019, for the first time, the call included data literacy as a topic. BiCDaS was involved in the call as well as in the selection process. The large number of data-literacy-related proposals and the highly innovative and promising concepts proposed showed that there is a strong demand among lecturers to tackle this topic.

B - 1.5 Lecture Series Data Science

As part of BiCDaS’s mission to increase data awareness and to support data science education, a lecture series on topics of data science was organised in the winter terms of 2017/18 and 2018/19. Internal and external high-profile speakers gave six lectures per term. True to our dedication to the idea of open science, we asked the speakers’ permission to record their respective lectures. These recordings were then post-processed⁴ and are now available on the BiCDaS website⁵.

B - 1.6 Data Science Day

A lecture series is an excellent method of advancing data awareness and education for persons who have already been exposed to the topic. Others will usually not easily accept its relevance and will, therefore, not be inclined to attend. To spread data awareness throughout the university, in May 2018, BiCDaS organised the first data science day at Bielefeld University. For the lecture series, people were invited. For the data science day, we aimed to bring the topic to where the people are.

Bielefeld University has a unique architectural layout. All parts of the university main building are connected through a main hall. This long-stretched corridor featuring gastronomy and shops is a common meeting place that most affiliates of the university pass through several times each day.

The data science day has been an exposition in the university main hall where twenty research groups from all over the university presented how they work with data. Expositions included manifold poster presentations, tool demonstrations, a guided tour through the GPU cluster of high-energy physics and many more. Exhibitors ranged from history (digital humanities) to biotechnology, from the university library to the data science master degree programme.

B - 1.7 Co-Organisation of Scientific Events

After the resounding success of the ‘Big Data made in Germany conference’ in 2017 in Berlin, of which BiCDaS has been a co-organiser, we continued this line of activity in 2018 by co-organising the 13th CeBiTec Symposium in Bielefeld under the topic ‘BigData in Medicine and Biotechnology’. Aside from providing financial support, BiCDaS particularly organised a panel discussion on ‘Legal and Ethical Aspects of Big Data in Medicine and Biotechnology’. Four experts from law studies and ethics participated in it and it marked the successful ending of the symposium and was well received.

Towards the end of 2018, we were approached by the Faculty of Business Administration and Economics who were considering to organise a workshop on ‘Data Science in Operations Management and Retailing’. BiCDaS gladly agreed to get involved and it became a very successful event, which attracted much attention from the local business community.

BiCDaS is also a sponsor of a summer school on time series analysis organised by the chair for econometrics at Bielefeld University⁶. Moreover, BiCDaS is an organising partner of a summer school on explainable data science by EuADS⁷ (see Section ‘B - 1.10 International Networking in Science - EuADS’). Both summer schools will be held in September 2019.

⁴ Using a special software for video recordings of educational events – Panopto → panopto.com

⁵ See https://uni-bielefeld.de/datascience/more/lecture_series.html (2 speakers disagreed to the recording.)

⁶ See <http://www.wiwi.uni-bielefeld.de/lehrbereiche/statoekoinf/oeko/mtsa>.

⁷ See <https://euads.org/summer-school-2019/>

B - 1.8 Support with Funding Proposals

In 2018, BiCDaS was involved in two funding proposals in two very different, but highly competitive and prestigious calls, both with a strong data science relation.

B - 1.8.1 Data Literacy Education

In April 2018, the Stifterverband für die Deutschen Wissenschaft published a call for proposals in data literacy education. At that time at Bielefeld University, many lecturers had already taken up this task. Their excellent and high-level teaching were seminal, but in the end isolated offers. BiCDaS and ZLL drafted a proposal to participate in this call, and while it was not selected from among the highly competitive pool of proposals, the protagonists at Bielefeld University felt that it contained many interesting ideas and concepts and, if implemented, might be a huge step forward for the education offered. Therefore, it has been funded from other sources (see Section 'B - 1.1 Data Science Teaching').

In June 2019 a new call for proposals has been published by the federal state of North Rhine-Westphalia. For this call, Bielefeld University has joint forces with the University of Paderborn to participate in that call with a further advanced and more concrete concept.

B - 1.8.2 Leibniz Science Campus

Over years, Bielefeld University and the Socio-Economic Panel Study (SOEP - Berlin) have been developing fruitful and involved collaborations. Those relations manifested in a connection-professorship and further blossomed when a BiCDaS member was called to the SOEP board.

In 2018, the Wissenschaftsgemeinschaft Gottfried-Wilhelm Leibniz (WGL)⁸ published a call for proposals for so-called Leibniz-Science-Campuses, i.e., collaborations of Leibniz Institutes with universities. SOEP (a Leibniz Institute) and Bielefeld University applied with a proposal aiming to enrich SOEP data with geospatial features. Data science is a red thread throughout the proposal and BiCDaS has been an active partner during the conceptualisation and writing.

Starting from two highly motivating reviews that unfortunately resulted in a rejection of the proposal, the partners are currently preparing a resubmission based on the reviewers' feedback.

B - 1.9 Research Data Management

Bielefeld University has been an early adopter of research data management as a strategic topic. After issuing a research data management policy in 2013, the topic had continuously been pursued at Bielefeld University in different forms.

The demand for support on this topic has continuously increased among researchers, both in the amount and in the spectrum of the services requested. To adapt, the Bielefeld IT Service Centre (BITS), the University Library and BiCDaS joined forces in early 2018 to conceptualise a competence centre for research data management as the next step in that development. After a positive internal evaluation, the competence centre for research data management was inaugurated in November 2018 (see Section 'A - 1.4.6 Competence Centre for Research Data Management').

B - 1.10 International Networking in Science - EuADS

In 2018, Bielefeld University applied, on the initiative of BiCDaS, for membership in the European Association for Data Science (EuADS). Since then BiCDaS has become involved in the EuADS administrative office as well as in the EuADS board and, since November 2018, in its presidium. The work for EuADS offers ample opportunities for networking and to increase BiCDaS's international

⁸ See <https://www.leibniz-gemeinschaft.de/>

visibility, e.g., in the EuADS summer school on explainable data science (see Section ‘B - 1.7 Co-Organisation of Scientific Events’) or in the development of a 2-sided platform for data science jobs.

B - 1.11 Exchange with Companies

The topic of data science in general and BiCDaS in particular has attracted much attention from companies in the region of East Westfalia Lippe. We are receiving requests for cooperation from various fields. While we appreciate such interest, any cooperation needs to feature a clear research focus and allow us to maintain our independence.

B - 1.12 Exchange with Faculty of Medicine

Bielefeld University is currently establishing a faculty of medicine. Medicine is nowadays a data-heavy science facing unique challenges such as the need for data representations allowing for quick decisions, a high reliability of results and a high degree of data sensitivity. Therefore, the Faculty of Medicine has a high interest in data science topics. Additionally, or maybe consequentially, data science has been defined to be a central topic in research and teaching of the new faculty. BiCDaS initiated an exchange with those responsible for developing the research profile, curriculum and (data) infrastructure. Together, we are identifying possibilities where BiCDaS can contribute to the development and are defining BiCDaS’s role in the future faculty.

B - 1.13 Support of GPU-HPC cluster

While Bielefeld University has access to CPU-based HPC capacity at the PC² at the University of Paderborn, there has until recently been no easy way to access GPU-based (i.e., highly parallel) HPC, except for the Faculty of Physics, which has maintained and used such capacity for a number of years now. Owing to its funding, this capacity was earmarked for use in physics only.

With the support of the rectorate, an existing funding for the modernisation of the physics GPU cluster could be extended. The new cluster became operational in May 2019, and a part of its computing capacity is now for general access by researchers from Bielefeld University. BiCDaS has accompanied this process and we are currently considering the formation of an interest group on GPU computing (see Section ‘B - 3.1 Interdisciplinary Dialog (DataLabs)’).

B - 1.14 Coordinating Activities in Context with NFDI

On the initiative of the German government, a national research data infrastructure (NFDI) is currently being developed across Germany. This infrastructure will be an interconnected effort and driven by the research community. Currently, consortia of researchers from all fields of science are developing concepts for a field-specific, interconnected research data infrastructure.



Figure 2 - An example for the BiCDaS news ticker.

BiCDaS sought to coordinate and connect the activities taking place in that regard within Bielefeld University and, therefore, invited interested researchers to a NFDI round table. The approximately 20 researchers who attended, considered the two meetings (up to now) as extremely fruitful.

B - 2 Communication

B - 2.1 Internal Communication

Regarding internal communication of news around BiCDaS, the round table is the main venue for sharing information. As part of the consideration on how the amount of benefits for BiCDaS researchers might be increased, BiCDaS started a biweekly news ticker service in November 2018. Each news item is summarised in two sentences and linked to the news' source.

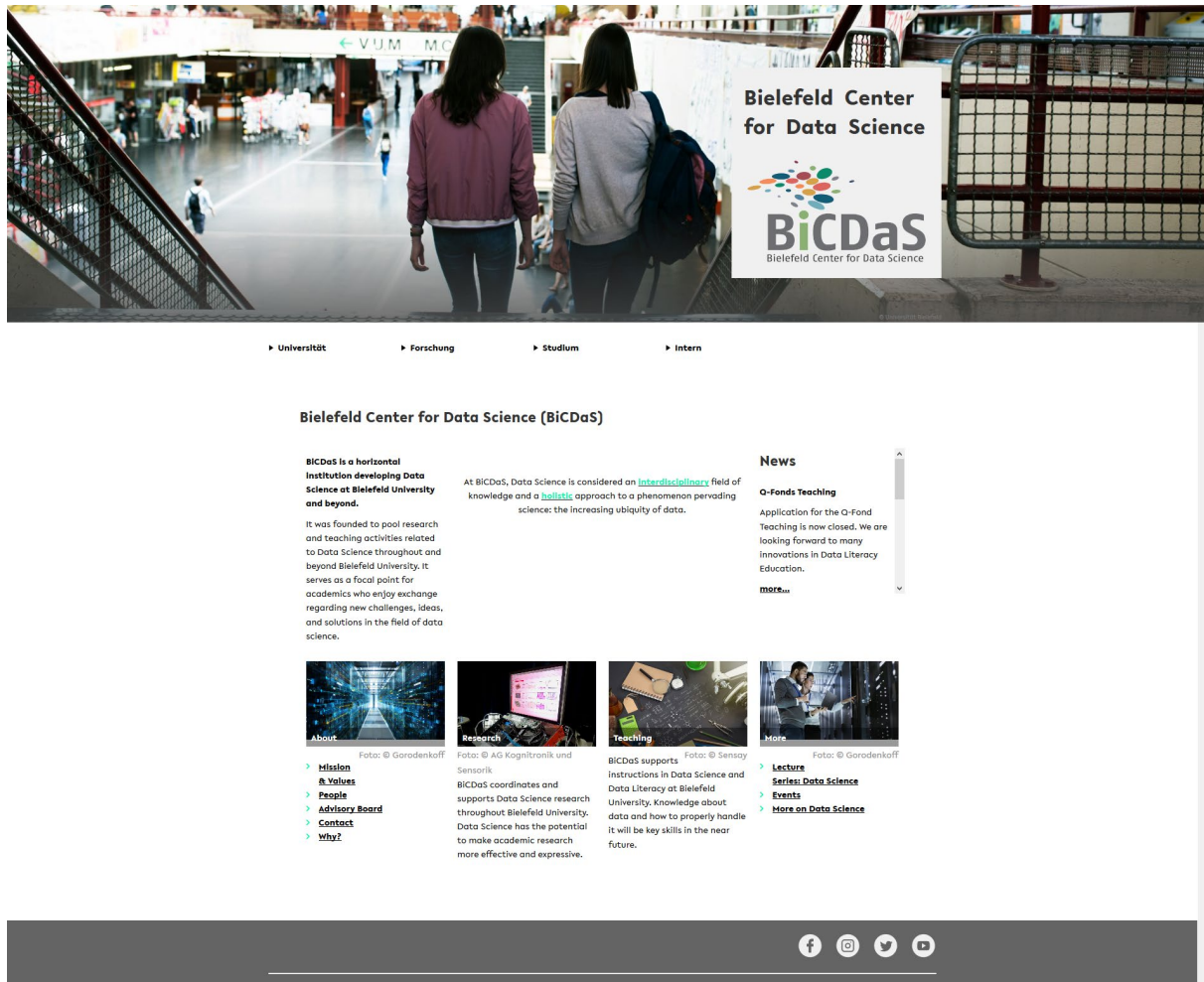


Figure 3 - BiCDaS Landing Page (taken June 18th 2019)

B - 2.2 External Communication

An appealing web presence is nowadays the central point for external communication. External stakeholders, including reviewers of funding proposals, can be expected to inspect the web presence as one source of information. Therefore, BiCDaS was quick to first have an informative web presence after it was founded. For the same reasons, we decided to redesign the website in 2018 with some additional effort.⁹ This web presence also features a selection of all data-science-related publications and all research data publications by researchers from Bielefeld University.

A second avenue of digital external communication used at BiCDaS is Twitter. Nowadays, Twitter has become a major method of networking for researchers. BiCDaS uses Twitter actively to increase its

⁹ See <https://www.uni-bielefeld.de/datascience/>

international visibility, to maintain contact with partners and researchers and as a means of public relation. BiCDaS typically produces two to three tweets a week. At the time of this writing, BiCDaS has about 570 followers and the number is constantly increasing.

B - 3 Future Plans/Strategies

BiCDaS is meant to generate a net benefit for the researchers at Bielefeld University. If BiCDaS does not make their work easier and more effective in the long run, it would have failed in its mission. Hence, in summer 2018, we took the time to study thoroughly some strategic considerations and how we can achieve this goal more effectively.

B - 3.1 Interdisciplinary Dialog (DataLabs)

One way by which we want to foster interdisciplinarity is to initiate scientific dialogs that are not structured along the lines of scientific disciplines, but rather along the lines of data that scientists frequently work with. In late 2018, we started to concretise these ideas, identifying four major types of scientific data (image data, text, graph data and time series) that we think will cover almost all of the research data at Bielefeld University.

Simultaneously, we conceptualised what we call DataLabs. Those are meant as an agile format of (internal) research cooperation. Furthermore, these DataLabs shall be marked by topic centricity and a distinct need orientation. Whenever research interests meet these criteria, a DataLab can be created within a short period of time – centred along either (i) a specific data infrastructure or tool, (ii) a specific, versatile data set, or (iii) a set of methods and/or models.

At the time of this writing, we were planning to organise informal meetings as incubators for such DataLabs. These meetings will pursue two main goals: (i) the identification by researchers of other research projects that use the same basic data type, and (ii) the formation of new DataLabs with their outline being determined by the researchers themselves.

B - 3.2 Stimulate Funding Proposals

Data science is a very broad and ubiquitous topic. For that reason, we have chosen the form of a horizontal institution for BiCDaS traversing all faculties. This also means that BiCDaS has an excellent position to stimulate interdisciplinary funding proposals. BiCDaS can bring together the relevant researchers, and data science can serve as a red thread throughout the proposed concept.

Therefore, it has been decided that the BiCDaS executive director will actively search, together with the Department for Research Administration and Technology Transfer, for calls that fit the profile of BiCDaS researchers and stimulates proposals.

B - 3.3 Using Momentum

Data science is a trending topic in society, economy and science. It is heavily covered in the media, and ample funding opportunities have been created in this field in recent years. Additionally, Bielefeld University has made data science a central strategic topic in many aspects. BiCDaS aims to use this momentum for the benefit of our researchers and lecturers.

Part C – Conclusion

BiCDaS is a decentralised, horizontal institution aiming to improve data work throughout Bielefeld University. As a rather young institution, it underwent a phase of extensive networking and conceptualisation, which it is currently leaving behind. We have great hopes that the first meeting of the BiCDaS advisory board in September 2019 will provide us with valuable feedback on these developments, concepts and future plans and that BiCDaS will thrive as a productive and integrative force across Bielefeld University.