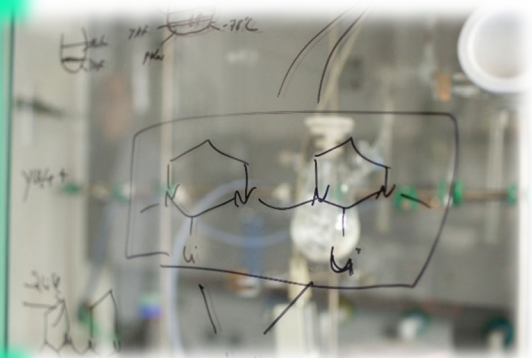


Consolidation and Progress

BiCDaS Activity Report – 2019 - 2021



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, Juliane Theiß, Katharina Weiß und Reinhold Decker
bicdas@uni-bielefeld.de

Content

- Part A BiCDaS Portrait 1
 - A - 1 BiCDaS Goals and Ideas 1
 - A - 2 Structure 2
 - A - 2.1 Round Table Data Science (BiCDaS General Assembly) 2
 - A - 2.2 Staff Members..... 2
 - A - 2.3 Scientific Advisory Board..... 3
 - A - 3 Our Academic Context..... 3
- Part B BiCDaS from 06-2019 to 06-2021 5
 - B - 1 Activities 5
 - B - 1.1 Data Science Teaching..... 6
 - B - 1.2 Scientific Events..... 7
 - B - 1.3 Support with Funding Proposals 8
 - B - 1.4 Internal Exchange 9
 - B - 1.5 Developing BiCDaS 10
 - B - 1.6 Shaping Data Science 11
 - B - 1.7 Outreach..... 12
 - B - 1.8 Miscellaneous..... 12
 - B - 2 Communication 13
 - B - 2.1 Internal Communication 13
 - B - 2.2 External Communication 13
 - B - 3 Future Plans/Strategies 13
 - B - 4 BiCDaS and SARS-Cov-2..... 15
- Part C – Conclusion..... 15

Executive Summary

The **Bielefeld Center for Data Science (BiCDaS)** is a highly interdisciplinary network at **Bielefeld University** that synergises researchers from all faculties and several central institutions. It currently has two full-time positions and 59 members from all parts of Bielefeld University.

The activities from June 2019 to June 2021 (the report period) encompass:

Developing **data science teaching**, especially the DaLiS@OWL project on data literacy, encompassing the conceptualisation and implementation of diverse courses, a data literacy module, a data literacy competence framework and preparations for a data literacy certificate to be obtained by graduates.

(Co-)organising and sponsoring **scientific events**, particularly three different summer schools and a lecture series, all advancing the topic of data science.

Supporting **funding proposals** in many different roles, from the solely responsible application and implementation on the one hand to the role as proof of the data science competence of Bielefeld University in applications from third parties on the other hand.

Developing cooperation and **outreach** with societal actors, such as foundations or associations, Bielefeld municipality and enterprises. Mutual benefit and scientific independence are the cornerstone of our activities in that regard.

Further **developing BiCDaS**, compelled by the continual development of the field of data science, which makes a static institution infeasible for the time being. Bringing to bear its unique perspective, BiCDaS is also engaged in developing **Data Science** in Germany and Europe.

BiCDaS has recently faced different challenges, particularly due to the ongoing pandemic. It is also planned to perpetuate BiCDaS by the end of the year, and preparations for this are another recent key activity.

In total, BiCDaS has found its role at the university and is about to flourish, maximising its utility for researchers and scholars, and the university in general.

Part A BiCDaS Portrait

Here, we discuss the major factors defining BiCDaS: (i) our mission and approach to data science, (ii) our vision for BiCDaS and Data Science at Bielefeld University, (iii) the context in which BiCDaS resides and (iv) BiCDaS's structure.

A - 1 BiCDaS Goals and Ideas

In one sentence, we summarise BiCDaS's mission as:

BiCDaS's Mission

*Unlocking data potentials with scientists, lecturers
and students throughout Bielefeld University*

It is, hence, our belief that data harbour great potential for Bielefeld University, above the considerable potential already explored/exploited. This potential lies within the handling and analysis of data, naturally, but even more so in the integrative, trans-disciplinary and transformative force, data can be within an academic context.

Furthermore, we are convinced that realising this potential to the maximum possible amount cannot be a stand-alone mission. Rather, we are seeking exchange – within the university and beyond it – with partners from science, politics and economy.

The activities of BiCDaS are structured into five key areas:

Key Areas of Activity

Supporting data science research

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

Advertising data science/data awareness

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

Supporting data science in academic research

BiCDaS supports academics in incorporating data science into their daily work. This is achieved, for instance, by providing data literacy training or by organising 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 on data science and enshrine data literacy as general objectives in academic education.

Data science as a bridge to the world

We believe that data science can serve as a connecting element/as an integrative force within the university and beyond. Data can be a common denominator for cooperation with, e.g. economic enterprises and/or social associations.

Besides these five key areas of activity, two horizontal values/topics guide BiCDaS's work.

Horizontal Topics

Data-Centred Scientific Dialogue

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 disciplinary boundaries are consequentially upstaged. Interdisciplinarity easily initiates 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 omni-lateral, 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, on algorithm usage and accountability and on how the emerging 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.

A - 2 Structure

BiCDaS has 59 members from all faculties and different central institutions. Most of them are researchers, complemented by some non-scientific staff, who are concerned with data science in their daily work (e.g. from the university library or BITS).

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

The round table data science is where BiCDaS originated, and nowadays, it is the 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 12 times. The original purpose of the round table was to synergise researchers and employees of Bielefeld University with an interest in data science and to discuss how this topic is best ploughed at Bielefeld University. Forming the Bielefeld Centre for Data Science was the answer given by the round table.

BiCDaS Origin

After an executive director was appointed, the round table played the role of a steering committee of the 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.

Role of the Round Table

A - 2.2 Staff Members

BiCDaS has two full-time positions: the executive director and the coordinator of ‘data literacy education’. BiCDaS does not incorporate any formal/legal entity up to now, and thus, its staff members are employed at the Department for Digital Transformation and Process Organisation. However, we intend to change that (see ‘Institutionalisation’ in Section B - 3 ‘Future Plans/Strategies’).

The position of executive director was created as a full-time position by the university rectorate after the round table data science decided to establish BiCDaS. The executive director’s task is to develop BiCDaS. For that, the position has been provided with a great deal of autonomy. The executive director reports to the round table and coordinates activities with the vice-rector for Information Infrastructure and Business Relations.

Executive Director

BiCDaS identifies data literacy education as a field of activity in which it can be highly effective. Data literacy means equipping students with data competencies that will enable them to leverage future jobs and participate in societal discourses, which are presumed to become increasingly data-driven (see Section B - 1.1 'Data Science Teaching'). The coordinator 'data literacy education' has advanced this project as a BiCDaS staff member and in close co-operation with ZLL and the programme 'richtig einsteigen.' (see ZLL in Section A - 3 'Our Academic Context'). Also, the coordinator helms the consortium project DataLiteracySkills@OWL (DaLiS@OWL, see sections B - 1.1 'Data Science Teaching' and B - 1.3 'Support with Funding Proposals') in which Bielefeld University collaborates with the University of Applied Sciences, Bielefeld, and Paderborn University.

Coordinator Data
Literacy Education

A - 2.3 Scientific Advisory Board

BiCDaS, as a horizontal institution, stretching across all faculties and involving different central institutions, faces unique challenges in the university environment. This same structure alongside the high popularity of the topic of data science offers unique possibilities for the benefit of BiCDaS members and the university in general.

To leverage these opportunities while facing the challenges adequately, it is vital to solicit the advice of external, highly experienced, academic professionals. We do this once a year on our advisory board, currently comprising six highly skilled 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 'BiCDaS Goals and Ideas').

A - 3 Our Academic Context

Bielefeld University is a medium-sized university with 24,940 students, 1,895 academic staff and 1,220 technical and administrative staff¹. It is currently involved in five Collaborative Research Centres (CRC-TRR 211, CRC-TRR 212, CRC 1283, CRC 1288 and CRC-TRR 318), houses 15 graduate schools and offers 51 courses on bachelor- and 60 courses on master-level alongside 12 PhD courses. It has a spatially compact campus with all institutions and faculties assembled in one place.

Bielefeld University

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 in achieving BiCDaS's goals.

Furthermore, the increasing significance of research data management was recognised by Bielefeld University at an early stage. Hence, it was among the first German universities to issue a policy on research data management, and it has continually pursued this topic ever since. Thus, Bielefeld University is presently spearheading 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. Subsequently, data science in medicine has been identified as a central topic in education and research.

¹ According to the Annual Statistical Report of Bielefeld University – year 2019.

BiCDaS is highly interconnected across Bielefeld University. We will briefly introduce some of our academic partners and institutions in alphabetical order (by their English names).

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

BITS

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 is, today, exemplary in its research data management, computing infrastructure and research software.

CeBiTec

At CeBiTec, a research cloud providing software applications for various analytic tasks powered by HPC resources and extensive cloud storage has been developed. The cloud services at CeBiTec are included in the German Network for Bioinformatics (de.NBI). The de.NBI cloud is being administered from CeBiTec, and the current de.NBI chair is a BiCDaS member. CeBiTec is a major player in different proposals towards the emerging national research data infrastructure (NFDI) and has, in early 2020, received an extension of its funding for two years, allowing it to lead this ongoing discussion to its logical conclusion.

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 study programme on statistical sciences alongside a master programme on data science (see Section B - 1.1 'Data Science Teaching'). ZeSt also hosts the Statistical Consulting Centre (Statistisches Beratungszentrum – StaBeCe), which offers advice on matters of statistics throughout the data life cycle.

ZeSt

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 advertising positions, traineeships or joint theses, which happens regularly.

ZLL

ZLL also featured 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 at Bielefeld University (see Sections B - 1.1 'Data Science Teaching' and B - 1.3 'Support with Funding Proposals'). The programme ended in March 2021.

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 also part of CITEC.

CITEC

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, 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 of increasing awareness of the importance of research data management throughout the university. Its funding has meanwhile been perpetuated.

Competence Centre
for Research Data
Management

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

The Research Centre for Mathematical Modelling (RCM²) focuses on the application of mathematical methods within the sciences. Currently, it focuses on mathematical biosciences, mathematical physics and applied dynamical systems. In the area of data science, RCM² investigates model-based approaches and, thus, contributes to 'explainable' data science. Such approaches currently play a decisive role in the evaluation of 'big data' in social networks and neural, genetic and metabolic systems, etc.

RCM²

The DFG-funded interdisciplinary collaborative research centre 'Practices of comparing. Ordering and changing the world' investigates the role of the societal practice of comparing from different scientific perspectives: sociology, anthropology, history, literature studies, political sciences, law and others.

CRC 1288

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

The University Library in Bielefeld nowadays is a 21st century general provider of information services. Besides standard library services, it offers numerous digital services. As such, the Publications at Bielefeld University (PUB)² system is hosted and developed at the library, which is a catalogue containing classical paper publications alongside research data publications across domains. The PUB system is a central element of the open-access strategy of Bielefeld University.³ Furthermore, the library's IT team hosts different open-source software programmes, such as a Gitlab or a Share Latex instance.

University Library

Part B BiCDaS from 06-2019 to 06-2021

Here, we will briefly report on BiCDaS and its activities from June 2019 to June 2021 and some future plans and strategic considerations.

B - 1 Activities

The years 2019–2021 were a time of consolidation and evaluation for BiCDaS. The efficiency of the activities we developed during the starting phase in 2017 and 2018 required assessment. Also, we prepared for establishing a formal structure, which we had neglected during the starting phase, favouring agile development.

² See <https://pub.uni-bielefeld.de>

³ See <https://oa.uni-bielefeld.de/index.html> (German)

B - 1.1 Data Science Teaching

Teaching in the field of data science is one of the five key areas of activity of BiCDaS (see Section A - 1 'BiCDaS Goals and Ideas'). At Bielefeld University, this is primarily done through a data science master programme, which is insofar exceptional, as it strongly emphasises 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 RCM² in Section A - 3 'Our Academic Context'). 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 these with its equally extensive expertise in algorithms and machine learning, as equal partners.

Data Science Master

With the data science master programme well on the way, BiCDaS could concentrate on the second branch of data-centred teaching at Bielefeld University – data literacy education (DLE). Data literacy aims at integrating data competencies into the curricula of all disciplines, and in 2018, BiCDaS initiated a project aiming at that goal. The project is governed by a multi-faceted team: lecturers from different career stages, delegates of the rectorate, ZLL (compare 'ZLL' in Section A - 3 'Our Academic Context') and BiCDaS, alongside student representatives and stakeholders from local business enterprises and society.

Data Literacy Education

The team of 'richtig einsteigen.' (see 'ZLL' in Section A - 3 'Our Academic Context') already had extensive experience in incorporating literal (LitKom) and mathematical (MatKom) competencies in curricula across study programmes. To leverage this experience and existing infrastructure, data literacy was implemented as an additional pillar (DatKom) within 'richtig einsteigen.'

'Richtig einsteigen.' ended in March 2021. Already in March 2020, the consortium project DataLiteracySkills@OWL⁴ (DaLiS@OWL) started. The University of Applied Sciences (FH), Bielefeld, Paderborn University and Bielefeld University participate in this consortium. An interdisciplinary dialogue among lecturers involved in data literacy education was initiated and supported. In addition, measures were described to increase data awareness throughout the university.

DaLiS@OWL

Initially, the data literacy project team, together with lecturers from several faculties, initiated the development of a sustained data literacy module, which is open to all undergraduate students starting in October 2019. Students can individually choose from different courses offered by different faculties to expand their knowledge, acquire additional data competencies, and complete the data literacy module. This portfolio is continually being expanded.

Data Literacy Module

On that note, in Q2-2021, BiCDaS received a grant for developing a new interdisciplinary hands-on course in data storytelling. The project Bielefeld Data Narratives (Bi_DaNte) will enrich the portfolio of courses available to students in the data literacy module.

Bi_DaNte

In the winter term 2019/2020, lecturers could submit proposals for innovative teaching ideas to the quality fund teaching. For the second time, data literacy was one of the topics. Many great ideas were submitted, and many of them touched upon data literacy. Ten ideas could be funded, many of which are also interesting for the Data Literacy module.

Call for DL Teaching

⁴ The region in which the consortium partners are located is called Ostwestfalen-Lippe (OWL)

The central part of the module is, however, the interdisciplinary lecture series ‘Big data is watching you! How to deal with data in the modern world’, which was launched in October 2019. In 15 lessons, nine lecturers from Bielefeld University and two external guests brought basic data skills to students from 42 different major subjects. Students obtained an overview of the role of data in science, economy and society, data analysis and the transformation of data into knowledge. In two embedded tutorials, the students were also equipped with basic skills in R.

Lecture Series DLE

Based on student feedback, the content and mode of teaching of the lecture series were revised for the second iteration starting from November 2020. In that process, the lecture series was renamed: ‘Forschen, Handeln und Gestalten in der Welt der Daten’ [Investigating, acting and creating in the world of data].⁵ Due to the concurrent SARS-CoV-2 pandemic, the lecture series has been further adapted, aiming to accommodate the needs of remote teaching and asynchronous learning.

The partners in the DaLiS@OWL project have agreed to develop a data literacy framework, a catalogue of learning goals to be achieved by students to become data literate graduates. This framework was developed under the aegis of the University of Paderborn and was presented to the partners in early 2021 with the request for comments. It is based on different others such frameworks, but adapts these catalogues to the local requirements – which is why local enterprises and society have been consulted in the process.

Data Literacy Framework

In the future, students who have fulfilled these requirements can obtain a DaLiS@OWL certificate. This certificate will be established as a ‘brand’ for local companies looking for data-competent applicants, giving these graduates a decisive advantage.

Data Literacy Certificate

B - 1.2 Scientific Events

(Co-)organisation of scientific events allows BiCDaS to set impulses and strengthen its scientific profile. In some cases, BiCDaS also acts as a sponsor for such events. Since much of this report’s time frame falls into the SARS-CoV-2 pandemic, these activities have been unusually highly demanding.

As part of BiCDaS’s mission to increase data awareness and to support data science education, each winter term, a lecture series featuring six high-profile speakers was organised. True to our dedication to the idea of open science, we asked for the speakers’ permission to record their respective lectures⁵. In the winter term 2019/20, the third edition of the lecture series data science occurred. This was concluded by a talk by Prof Douglas Bates (U Wisconsin, Madison, US), who is a highly recognised statistician and a core developer of the data science languages R and Julia.

Lecture Series DS

To optimise our offers to the researches, we are currently considering transferring the lecture series to a less formal format by offering lectures from high-profile speakers whenever the opportunity presents itself while using this as a vehicle to set impulses.

BiCDaS members are engaged in the European Association for Data Science (EuADS), a network of high-profile researchers, as members of the board and presidium (see Section B - 1.6 ‘Shaping Data Science’). In September 2019, the EuADS organised an international summer school on the topic of ‘Explainable Data Science’.⁶ BiCDaS has been a sponsor of

EuADS Summer School

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

⁶ See <https://euads.org/summer-school-2019/>

this successful event, featuring highly prestigious researchers such as Christopher Bishop (machine learning pioneer, today leading Microsoft research in Cambridge, UK).

After its success in 2019, we aim to hold another EuADS Summer School on the topic 'Data Science for Social Media'. The summer school features high-profile speakers from the field. The event was originally scheduled for 2020 but had to be postponed repeatedly due to the pandemic. On July 19th, 2021, EuADS will also host an online symposium titled 'Data, Social Media, and Society,' featuring Alex Pentland of the MIT Media Lab.

BiCDaS also sponsored the summer school 'Modern Topics on Time Series Analysis', which was organised by Bielefeld University's chair of econometrics. The event attracted an international audience featuring some high-profile speakers, such as Helmut Lütkepohl.

Summer School MTTSA

B - 1.3 Support with Funding Proposals

BiCDaS was involved in different funding proposals in various ways. We have learnt to appreciate our extraordinary position beyond disciplinary research, allowing us to provide a neutral framework, particularly for interdisciplinary proposals.

In June 2019, a call for proposals was published by the federal state of North Rhine-Westphalia on the topic of data literacy education. Bielefeld University, led by BiCDaS and ZLL, was already highly active in this field at the time (compare BiCDaS Activity Report 2018/19 'Phases of Development'), and we leveraged this to intensify our activities and spread the idea of data literacy education beyond Bielefeld University. Hence, Bielefeld University collaborated with Paderborn University and the University of Applied Sciences (FH) Bielefeld on a proposal for a joint project called DataLiteracySkills@OWL⁷. This proposal has been successful, and the project was funded with approximately € 900,000 for three years, which started 1st of March 2020.

DaLiS@OWL

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

SOEP@UniBi

The Wissenschaftsgemeinschaft Gottfried-Wilhelm Leibniz (WGL)⁸ issued a call for proposals in 2018 for so-called Leibniz-Science-Campuses, i.e. collaborations of Leibniz-Institutes with universities. Bielefeld University and SOEP (a Leibniz Institute) applied a proposal aiming to enrich SOEP data with geospatial features. Data science was a red thread throughout the proposal and BiCDaS has been instrumental during the conceptualisation and writing. In 2019, in the second attempt, this proposal was successful, and the SOEP@UniBi Leibniz Science Campus started in November 2020.

BiCDaS remains involved in the project, as BiCDaS executive director is a Co-PI for the subproject 4. This is an infrastructure subproject addressing many concerns regarding research data and the extraction of regional indicators from unstructured data. It thereby lays the basis for the work of the other three subprojects.

Participation in the Leibniz Science Campus

A particular challenge is how to establish an SEOP training dataset, which will be the centre piece of the SOEP@UniBi region hub. This dataset will be a 50% sample of the SOEP data with resolution down to the 'Raumordnungsebene', i.e. high-resolution survey data

⁷ Ostwestfalen-Lippe (OWL) being the region in which all three partners subside and also the context in which a variety of joint activities in various fields already take place.

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

containing highly personal information. There is no precedent for such high-resolution data being hosted outside DIW in Berlin, where SOEP resides. The first central task of the subproject 4 is, thus, to establish the prerequisites for this transfer, technical or organisational, and regarding data protection/security. The executive director of BiCDaS drives these activities.

The project Bielefeld Data Narratives (Bi_DaNte) has been proposed to the German KI-Campus⁹ initiative and will be funded starting summer 2021. The aim is to develop a course that will equip students with practical skills in data storytelling, i.e. the audience tailored communication of analytic results. The project will be funded for seven months, after which the seminar will be perpetuated. The seminar will be eligible for the Data Literacy Module (see 'Data Literacy Module' in Section B - 1.1 'Data Science Teaching').

Bi_DaNte

Furthermore, BiCDaS has set impulses, provided expertise or has been placated in different funding proposals to document the expertise in data science at Bielefeld University.

B - 1.4 Internal Exchange

Data science can be an integrative force in academia, boosting interdisciplinarity and collaboration. BiCDaS harnesses this potential through various data-centred exchange formats.

BiCDaS has conceptualised a new, agile, internal cooperation format, called DataLabs. The aim of DataLabs is twofold: (i) fostering exchange among researchers and (ii) bundling expertise on a data-related topic in an externally visible manner. The DataLabs also help in reacting quickly to funding calls in their respective fields. This is particularly valuable since the time between a call and the respective deadline has become increasingly tight in recent years.

Data Labs

In late 2019 and early 2020, two first DataLabs were formed: the DataLab on High-Performance-Computing and Advanced Simulation formed around the GPU-Cluster inaugurated in May 2019. It connects researchers from economics and physics. The DataLab on Digital Humanities joins different scholars from the humanities at Bielefeld University alongside an external project partner from the University of Leipzig. In 2021, a DataLab Corona – Data Impacts and Solutions – was formed, bundling the research at Bielefeld on the pressing contemporary matters emerging from the ongoing pandemic.

Bielefeld University has been active in different fields of Open Science for many years. Particularly in the field of Open Access and Open Data, Bielefeld University has been at the forefront of development in Germany. Currently, the University Library and BiCDaS are aiming to form an Open Science Network Bielefeld as an exchange platform for researchers. While anchoring the Open Science Network centrally will ensure continuity, the partners intend to engage researchers in any stage of the development, particularly during the conception of the Open Science Network.

Open Science Network

A kick-off meeting for the Open Science Network was held on April 27th. About 30 researchers from various disciplines attended the meeting and discussed different projects that the network might tackle. Prof. Felix Schönbrodt, head of the Open Science Center at LMU, Munich, gave an impulse talk preceding the discussion. Subsequently, two project teams were formed, one aiming to develop peer formats on open science methods and the

⁹ See <https://ki-campus.org/>

other developing an open science seminar that will be eligible in the Data Literacy Module (see ‘Data Literacy Module’ in Section B - 1.1 ‘Data Science Teaching’).

A national research data infrastructure (NFDI) is currently being developed following German government initiatives. Consortia of researchers from all fields of science and humanities are developing concepts for field-specific, interconnected research data infrastructures.

Coordination NFDI

BiCDaS coordinates the efforts to participate in these consortia, undertaken in various disciplines throughout the university, particularly through invitation of recurring meetings of a round table NFDI. The researchers found these meetings effective and helpful.

In October 2020, an association was founded as a centre piece of the governance of NFDI by the federal government and the states of Germany. After discussion at the (Bielefeld) round table NFDI and subsequent approval by the rectorate, Bielefeld University successfully applied for membership. BiCDaS has initiated and implemented this process.

In June 2019, BiCDaS advisory board member Prof Baillet (University of Le Man) presented on digital learning in the field of digital humanities in a so-called ‘Lehrbar Spezial’, organised by BiCDaS and ZLL (compare ‘ZLL’ in Section A - 3 ‘Our Academic Context’). The about twenty-five attendees were very satisfied with the event and engaged the speaker in a vivid discussion.

Lehrbar

BiCDaS organised another Lehrbar in early 2021: Our partners from Paderborn University have been invited to present on the data literacy framework they have been developing as part of the DaLiS@OWL project (see ‘Data Literacy Framework’ in Section B - 1.1 ‘Data Science Teaching’). This framework will, among others, form the basis for the planned OWL-Data Literacy Certificate, which students can obtain when graduating to prove their data skills to prospective employers.

The BiCDaS executive director became a board member of the Institute of Law in the Field of Intelligent Systems (Recht Intelligenter Systeme – RiT), thereby building a bridge and pursuing exchange between these two institutions while also exploring the relevance data science has for the law sciences (e.g. in legal tech or implications of model explainability).

Board RiT

B - 1.5 Developing BiCDaS

We are continuously developing ideas to further leverage BiCDaS for researchers and the university. We are taking feedback and inspiration from various sources: the dialogue with our members and the university’s administration, comparable institutions from around the world and the feedback provided by our advisory board. Based on the latter, in late 2019, we have been conducting a stakeholder-analysis of our activities so far, regarding their effectiveness, time and resources spent. In 2020, we focussed on adaptations of our activities and formats compelled by the pandemic. Based on the assumption that BiCDaS has found its role within the university by now, we intend to mirror the BiCDaS staff’s notion of that role with the researchers’ notion in 2021. Hence, we are conducting qualitative interviews with different BiCDaS members. Such yearly ‘strategy sprints’ have become a custom each summer break when there are fewer meetings, and most projects are on a vacation-induced hold.

Developing Concepts

The first session of the international scientific BiCDaS advisory board (see Section A - 2.3 ‘Scientific Advisory Board’) occurred on September 18th, 2019. Four of the five members of the advisory board were present at the four-hour meeting, which provided BiCDaS with

Advisory Board

valuable knowledge. The fifth member (Prof Baillot, U of Le Man), who could not attend on September 18th, had already visited BiCDaS in July and had given her feedback on BiCDaS's development and plans.

Meanwhile, a sixth member has joined the advisory board: Niko Beerenwinkel, Computational Biology, ETH Zurich.

Bielefeld University is currently establishing a faculty of medicine. Medicine is now a data-heavy science, facing unique challenges such as a need for data representations that allow for quick decisions, a high reliability of results and highly sensitive data. Thus, the Faculty of Medicine has a huge interest in data science topics. Also, or maybe consequentially, data science is a central topic in the research and teaching of the new faculty. The first professorial chairs were staffed, and Bielefeld University could recruit some excellent researchers with proven quantitative research profiles.

Exchange with Faculty of Medicine

While we are generally content with BiCDaS size (59 members currently), we are always aiming to maintain BiCDaS's interdisciplinary character. While we mostly accept applications for membership as they come, on some occasions, we actively recruit new members when certain disciplines/faculties are not or underrepresented. This has been the case for the faculty of chemistry and the faculty of psychology and sport sciences so far, both of which we could recruit new members from.

Recruiting Members

B - 1.6 Shaping Data Science

As an interdisciplinary centre at a major German research university, BiCDaS can offer a unique perspective on the role of data science in academia. We, therefore, weight in on the discussion and development on a national and international level.

National Level

We are engaged in the discussion on the topic on a national level via two leading scientific societies: The German 'Gesellschaft für Klassifikation' (GfKI) [Classification Society], also known as the Data Science Society, and the 'Gesellschaft für Informatik' (GI) [Informatics Society], precisely the Task Force on Data Science and Data Literacy at GI.

GfKI, GI

To synergise data literacy education protagonists from different universities and to foster the mutual exchange of ideas and concepts, the DATEV Stiftung Zukunft [DATEV Future Fund] and the Stifterverband für die Deutsche Wissenschaft [Donors' association for the promotion of humanities and sciences in Germany] have founded a national data literacy education network. Bielefeld University successfully applied for membership in this network. From October 2019, BiCDaS together with ZLL (compare 'ZLL' in Section A - 3 'Our Academic Context') represents Bielefeld University in the national data literacy education network. The regular exchange and peer consultations among the 14 networking institutions are valuable opportunities to partake in shaping data literacy in Germany.

DATEV Data Literacy Network

On the invitation of the organisers of the German Data Science Days at the LMU Munich, BiCDaS members had the opportunity to present their ideas about Data Science and Data Literacy Education in a panel discussion, concluding the two-day event.

German DS Days 2020

Inspired by German Data Science Days and backed by the research we had done in preparation for this event alongside our pre-existing expertise, we decided to expand this

Data Education Paper

research and transfer it into a paper publication.¹⁰ The paper project focus shifted, also thanks to the critical and productive insights from the reviewers, and the final paper's main contributions now lie in (a.) commenting on the role of diversity of educational contents in the field and (b.), more concretely, defining the term of 'data self-empowerment' as all skills required by an individual in an increasingly data-driven society. During this project, we also realised how disconnected the literature on data literacy is between Europe and the US. A phenomenon we partly attribute to the very distinct educational systems.

In the spring of 2020, the federal government of Germany prepared a federal data strategy. Part of these preparations were public consultations, particularly a public survey. BiCDaS's data literacy coordinator and executive director both took the chance to contribute to the process by participating in the survey. They could, thereby, also add their rare transdisciplinary, scientific, data-centred perspectives to the process.

Data Strategy Federal Government

International Level

In 2018, Bielefeld University applied on BiCDaS's initiative for membership in the European Association for Data Science (EuADS). Since then, BiCDaS has become involved in both the EuADS administrative office and board and, since November 2018, in the EuADS presidium. The work for EuADS offers ample opportunities for networking and to increase BiCDaS's international visibility, e.g. in the EuADS summer school on explainable data science (see 'EuADS Summer School' in Section B - 1.2 'Scientific Events'). Most importantly, BiCDaS can also, via EuADS, contribute to shaping the topic of data science on a European level.

EuADS

B - 1.7 Outreach

Data science can be an integrative force within academia and serve as a bridge to the outside world. We are actively pursuing scientific cooperation with partners from foundations, NGO, public administration and enterprises. The topic of data science in general and BiCDaS in particular has attracted much attention from companies in the region of Ostwestfalen-Lippe (OWL). We are receiving requests for cooperation from various fields. While we appreciate such interest, any prospective cooperation needs to feature a clear research focus, must be mutually beneficial and allow us to maintain our independence.

During the report period we have been pursuing cooperation, scientific dialogs, exchange and partnership with different foundations, enterprises, associations and municipalities. For more detailed information, please contact bicdas@uni-bielefeld.de!

B - 1.8 Miscellaneous

In early 2019, the Vice-Rector for Information Management asked BiCDaS to prepare an overview of the state of digitalisation at Bielefeld University. Trigger was a study by HIS-HE regarding the state of digitalisation at German universities.¹¹ At the end of the process in which BiCDaS could cash-in on its network structure and interviewed colleagues all across the university stood a presentation on the state of digitalisation, which has since been presented in various contexts and served as a basis for the next steps.

Digitalisation-Survey

¹⁰ Hachmeister, N., Theiß, J., Weiß, K., & Decker, R. (2021). Balancing Plurality and Educational Essence: Higher Education Between Data-Competent Professionals and Data Self-Empowered Citizens. *DATA*, 6(2), 1-15. <https://doi.org/10.3390/data6020010>

¹¹ H. Gilch et al (2019), *Digitalisierung der Hochschulen*, Expertenkommission Forschung und Innovation

B - 2 Communication

B - 2.1 Internal Communication

Regarding the internal communication of news around BiCDaS, the round table is the main venue for sharing information. Regarding news from the different fields of data science and the international data science community, BiCDaS started a biweekly news ticker service in November 2018. Each news item is summarised in two sentences and linked to the news' source.

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 has put considerable effort into an expressive web presence.¹² 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 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 940 followers and the number is constantly increasing.

B - 3 Future Plans/Strategies

In the past years, BiCDaS has defined its role within the university, spun a network of partners and collaborators within and beyond the university, and developed substantial traction within the university and the data science community via a rich and diverse portfolio of activities. Maintaining and increasing this traction while simultaneously undergoing the changes necessary to become a perpetual institution are the challenges that lie ahead.

BiCDaS is, up to now, not formally institutionalised. Being an informal entity united by the interest, commitment and good-will of the researchers allowed us for agile development. Institutionalisation

The screenshot displays the BiCDaS News Ticker interface. At the top, the BiCDaS logo is accompanied by the text 'News Ticker Woche 33 - 2019'. A navigation bar contains links for 'Forschung', 'Lehre', 'Politik', and 'Verschiedenes'. A vertical sidebar on the left lists various research areas: Biologie, Data Science, (Daten-)Ethik, Deep Learning, Digital Humanities, Digitalisierung, Erziehungswissenschaften, Forschungsdaten, Förderung, High-Performance-Computing, Künstliche Intelligenz, Lebenswissenschaften, Maschinelles Lernen, Mathematik, Medizin, Natural Language Processing, Open Science, Reproduzierbarkeit, Text Daten, and Veranstaltungen. The main content area features a news item titled 'Forschung' with the sub-header 'Digital Humanities + Maschinelles Lernen' and a date of '20.06.2019'. The text of the article discusses a Technology Review article about digital humanities and game theory. Below it, another news item is partially visible, titled 'Künstliche Intelligenz' with a date of '28.07.2019'.

An example for the BiCDaS news ticker.

¹² See <https://www.uni-bielefeld.de/bicdas/>

This is, however, not a viable form anymore, now that BiCDaS is about to be perpetuated. We are evaluating different options for a formal institution. Different choices also imply different roles of BiCDaS: from a high-profile research hub to a trans-disciplinary support structure. This is one of many reasons this decision should not be taken lightly. The



[► Universität](#)
[► Forschung](#)
[► Studium](#)
[► Intern](#)

Bielefeld Center for Data Science (BiCDaS)

News

- **Open Science Network gegründet**
Am 27. April fand ein Kick-off Treffen statt, an dem rund 30 Forschende teilgenommen haben.
 - **Journal paper on Data-related Education published**
Title: "Balancing Plurality and Educational Essence: Higher Education Between Data-Competent Professionals and Data Self-Empowered Citizens"
[more...](#)
 - **New DataLab**
On April 14th, 2021 the newly founded DataLab tackling interdisciplinary data topics around the Corona-Pandemic will meet the first time.
[more...](#)
- [More News](#)

BiCDaS is a horizontal institution developing Data Science at Bielefeld University and beyond.

It was founded to pool research and teaching activities related to Data Science throughout and beyond Bielefeld University. It serves as a focal point for academics who enjoy exchange regarding new challenges, ideas, and solutions in the field of data science.

»At BiCDaS, Data Science is considered an interdisciplinary field of knowledge and a holistic approach to a phenomenon pervading science: the increasing ubiquity of data.«

— BiCDaS Position Paper



About

- [Mission & Values](#)
- [People](#)
- [Advisory Board](#)
- [Activity Report 2019](#)
- [Contact](#)
- [Why?](#)

Research

BiCDaS coordinates and supports Data Science research throughout Bielefeld University. One central instrument are our new, agile, cross-faculty research collaborations, called **DataLabs**.

Teaching

BiCDaS supports instructions in Data Science and Data Literacy at Bielefeld University. Knowledge about data and how to properly handle it will be key skills in the near future.

More

- [Lecture Series: Data Science](#)
- [Events](#)
- [More on Data Science](#)



BiCDaS Landing Page (taken June 09th 2021)

institution traversing all faculties. This implies that BiCDaS is in an excellent position to stimulate interdisciplinary funding proposals. BiCDaS can unite the relevant researchers, and data science can serve as a red thread throughout the proposed concept. It has, therefore, 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 stimulate proposals.

perspective of the researchers is imperative in this process, as the benefits the researchers draw from BiCDaS ought to be maximised during the institutionalisation. Therefore, we have started qualitative interviews with different BiCDaS members on their perspectives on and wishes regarding BiCDaS. The first results indicate a clear division in affordances imposed on BiCDaS by researchers working on data science methods (e.g. statisticians, computer scientists and mathematicians) and researchers applying data science methods to their domains (e.g. natural sciences, social sciences and humanities).

The rectorate has mandated us to evaluate the different options, develop concepts and bring the institutionalisation of BiCDaS to a definite conclusion by the end of 2021.

Data science is a broader and ubiquitous topic, and BiCDaS is, thus, a horizontal

Stimulate Funding Proposals

B - 4 BiCDaS and SARS-Cov-2

Like virtually any other aspect of science or society in general, the ongoing SARS-CoV-2 pandemic has considerably affected BiCDaS' activities. BiCDaS is a horizontal institution and relies on its activities and role to a good extent on networking within in the university. Consequentially, what might have had the worst impact on BiCDaS' activities is the lack of opportunity for random encounters of colleagues and members on the campus. In retrospect, BiCDaS has, however, not been hit as hard as some by the pandemic.

While as soon as a standard video conference solution for Bielefeld University was in place (about April 2020, hence, pretty soon after the pandemic hit Germany), meetings with colleagues could be conducted effortlessly, we felt and feel that virtual meetings are and can only ever be an imperfect substitute. We therefore postponed many meetings in the (ultimately traitorous) hope that the pandemic might subside in summer 2020.

Even worse were the situations regarding conferences. While the German Data Science Days in Munich in February 2020 could be conducted as normal (compare 'German DS Days 2020' in Section B - 1.6 'Shaping Data Science'), basically all other conferences in 2020 had to be postponed. Particularly, the European Conference on Data Analysis (ECDA) and the EuADS Summer School had to be postponed/moved online throughout 2020 and 2021 (compare 'EuADS Summer School' in Section B - 1.2 'Scientific Events').

Another challenge of the pandemic lies in the visibility of our activities. For that, we have developed various activities in the digital sphere, e.g. the new webpage of the DaLiS@OWL project, virtual scientific talks and workshops.

SARS-CoV-2 is probably *the* hot topic in research in 2020/21. Hardly any other topic has seen an equally rapid growth in research output, funding, scientific discourse and public visibility. Even if the pandemic can be forced in retraction by several emerging vaccines, the pandemic, its consequences, potentials for stalling a next pandemic and doing better in many aspects when a next pandemic hits, will continue to concern researchers from various fields for years to come. Bielefeld researchers have been highly active in research on various aspects of the pandemic and in many disciplines. In early 2021, we therefore initiated the formation of a DataLab 'Corona – Data Impact and Solutions' fostering interdisciplinary exchange on the topic.

Part C – Conclusion

BiCDaS is a decentralised, horizontal institution aiming to improve data work throughout Bielefeld University. Currently, it is transitioning from an agile, commitment-driven, informal entity to a still agile, still commitment-driven, formal, perpetual entity. It faces many challenges stemming from this transition and from the ongoing pandemic. We are confidently tackling these challenges and aim to exploit the roles of data science as a driver of interdisciplinarity and of BiCDaS as a catalyst for new ideas and projects for the benefit of our members and the entire university.

Conduction of events and meetings

Visibility of BiCDaS Activities

DataLab Corona – Data Impacts and Solutions

Photo: © Universität Bielefeld