BERD@NFDI

A Data Marketplace to Foster Industry – Academia Collaboration

Frank Ohnesorge, University of Mannheim

March 2024
- 27 consortia represent a broad spectrum of scientific disciplines

- BERD@NFDI is a consortia of all business and economic related disciplines.

- BERD@NFDI strengthens the links with industry, as data from companies is playing an increasingly important role in research.
National Research Data Infrastructure (NFDI) – Objectives

Building
- networked information infrastructure & research data management

Establishment
- of standardized procedures and handling of research data

Development
- Of sustainable services and meta data standards

Connection
- of European & international platforms

Creation
- Reuse of existing data & integrity/quality/protection of data
Data as a common good for excellent research, organized by the scientific community in Germany.

Vision of NFDI
NFDI sections work on cross-disciplinary topics…

Training & Education

(Meta)data, Terminologies, Provenance

Common Infrastructures

Industry Engagement

Ethical, Legal and Social Aspects
BERD@NFDI builds a powerful platform for collecting, processing, analyzing and preserving Business, Economic and Related Data – all in one place.

- **Open**
  Linked unstructured and structured data

- **Fast and accessible computation**
  By cloud-based HPC solution

- **Best practices in ML**
  Platform provides guidance on methods

- **Reproducible and Transparent**
  Documented used data and methods

- **Management of the entire data life cycle**

We facilitate the integrated management of algorithms and data along the whole research cycle, with a special focus on unstructured (big) data such as video, image, audio, text or mobile data.

*Source: BERD@NFDI*
Beyond other services, BERD currently offers:

- **Open-Big Data Directory**
  Curated collection of high-quality, open-source datasets

- **BERD Marketplace**
  Bridging the gap between organizations and researchers to create value for both sides
The BERD@NFDI Data Marketplace bridges the gap between owners of unique and new types of data and researchers.
BERD@NFDI Data Marketplace supports organizations to balance...

Data-based value creation & Data-driven innovations

Data privacy & Data regulations
<table>
<thead>
<tr>
<th>Wharton Research Data Services</th>
<th>Chicago Booth KILTS Center Research Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="https://www.chicagobooth.edu/research/kilts/research-data" alt="Wharton Logo" /></td>
<td><img src="https://wrds-www.wharton.upenn.edu/" alt="Chicago Booth Logo" /></td>
</tr>
</tbody>
</table>

For **25+ years**, Wharton Research Data Services (WRDS) has supported users with targeted solutions that underpin research, reinforce learning, and enable discovery. WRDS advances comprehensive thought leadership by giving users the power to analyze complex information.

The Kilts Center is home to a rich archive of data available to academic researchers at Chicago Booth and beyond. Made possible through relationships with Booth alumni and corporate partners, these unique resources help academics around the world generate multi-disciplinary insights.

---

Close collaboration between academia and industry enables mutually beneficial relationships, producing high-quality, business-friendly articles and enabling access to data science talent and insights.

Source: https://www.chicagobooth.edu/research/kilts/research-data; https://wrds-www.wharton.upenn.edu/
Data challenge 2022: Product recommendations based on previous purchases

**Situation**
H&M Group is a family of brands and businesses with 53 online markets and approximately 4,850 stores. Their online store offers shoppers an extensive selection of products to browse through.

**Problem**
Overwhelmed customers do not find inspiration or what they are looking for and thus do not purchase.

**Objective**
Develop product recommendations based on data from previous transactions, as well as from customer and product meta data.

**Key outcomes**

- **Enhanced Experience**
  Personalized product recommendations simplify selection, reduce decision fatigue, and increase customer loyalty.

- **Increased Sales & Revenue**
  Targeted cross-selling and upselling through recommendations lead to higher conversion rates and increase the average order value.

- **Sustainability Impact**
  Reduced returns from informed recommendations result in reduced emissions and transportation, consistent with environmentally friendly practices.

Source: https://www.kaggle.com/competitions/h-and-m-personalized-fashion-recommendations
**Data challenge 2022: Predicting future loan defaults**

**Situation**
Modern life counts on the convenience of a credit card to make daily purchases. It saves us from carrying large amounts of cash and can also advance an entire purchase that can be paid for over an extended period.

**Problem**
Credit card issuers face risk of payback defaults with varying likelihood across the client portfolio.

**Objective**
Leverage an industrial scale data set to build a machine learning model that challenges the current model in production to predict credit default.

**Key outcomes**

- **Innovative risk assessment**
  Disrupt the credit default prediction model employed by the global leader in payment cards, fostering the development of cutting-edge solutions.

- **Enhanced customer journey and experience**
  Improved cardholder experience by streamlining credit card approval and ensuring a seamless and hassle-free process.

- **Pathways to success**
  Open doors to exciting career ventures at American Express, inviting participants to embark on an engaging and rewarding journey.

Source: [https://www.kaggle.com/competitions/amex-default-prediction](https://www.kaggle.com/competitions/amex-default-prediction)
Data challenge 2021: Single neuronal cell detection in microscopy images

Situation
Neurological disorders contribute significantly to global mortality and disability. Accurately segmenting neuronal cells in images, crucial for treatment evaluation, is a challenge that computer vision aims to overcome, enabling innovative drug discoveries.

Problem
Current methods to segment unique and irregular neuronal cells from images lack accuracy, hindering research progress.

Objective
Detect and delineate distinct objects of interest in biological images depicting neuronal cell types.

Key outcomes

- **Advances in Research**
  Successful results advance understanding of neurobiology via accurate segmentation of neuronal cell instances and data generation.

- **Enhanced Disease Insights**
  Researchers gain deeper insights into the effects of disease and treatments on neuronal cells, leading to more targeted therapies.

- **Innovative Drug Discoveries**
The breakthrough in accurate segmentation may facilitate the discovery of novel drugs, potentially reducing death and disability rates related to neurological disorders.

Source: https://www.kaggle.com/competitions/sartorius-cell-instance-segmentation
Recruitment opportunities
Collaboration with academia can serve as a platform for talent recruitment, identifying potential employees and partners.

Real-world impact
Real-world data and problems allow researchers to address practical questions that can be very motivating for participants.

Long-term relationships
Partnerships foster lasting relationships between companies and universities that lead to ongoing collaboration, knowledge sharing, and joint ventures.

Access to talent pool
Academia provides diverse talent, fresh perspectives, innovation, and advanced skills for complex challenges.

Research and techniques
Academia leads research. Partnerships offer firms cutting-edge technologies, methods and tools to gain a competitive edge.

Learning and development
Companies gain insights from academia on trends, methods, and emerging tech, fostering professional growth.

Innovation and creativity
Academic collaboration fosters innovative problem-solving with unconventional approaches and techniques.

Addressing major challenges
Joining impactful data challenges aligns a company with broader goals and shows social responsibility.

Source: BERD@NFDI
Offering data collaborations on BERD Data Marketplace

1. Visit BERD Platform

Visit the BERD platform and navigate towards “New Upload”

2. Login / create account

3. Define your “dataset / marketplace dataset”

Select resource type “dataset / marketplace dataset” and fill out information about your data offer to attract applicants. You can optionally add exemplary data snippets as files, but your original dataset is not uploaded. You later decide on operational access provision for successful applicants.

4. Submit data offer for review

Finally, submit your data collaboration offer by clicking on submit for review on the right-hand side.

5. Align with BERD team

A BERD team member will be in touch with you to consult with you and confirm / optimize your offer for the best possible outcome.

5. Start reviewing & approving applications

You remain in control through the entire process. You are the only one with access rights to review, approve or reject applications.

Your data remains secure, access is granted according to your preferences. Your data is never directly uploaded to BERD. You define with whom and how you like to share data.

You remain in power to define contractual obligations and policies with the users of your data. Agreements to your terms of use and license agreements can be signed with users.

Source: BERD@NFDI
1. Turning your companies’ data into innovation and business value by matching your company with leading academic research and data science teams in Germany, Europe and North America.

2. No risk of data getting into the wrong hands. You retain control of the data at all times, even during a collaboration.

3. Service is for free for companies as well as academic researchers, no hidden costs.
The BERD Data Marketplace is accessible NOW!

Visit us today:

[QR Code]

https://berd-platform.de/for-companies
Your key contacts

BERD & NFDI section “Industry Engagement”

Prof. Dr. Florian Stahl
University of Mannheim
florian.stahl@uni-mannheim.de

Isabella Hartig
University of Mannheim
isabella.hartig@uni-mannheim.de

Frank Ohnesorge
University of Mannheim
frank.ohnesorge@uni-mannheim.de

BERD Data Marketplace

Direct contacts

NFDI section “Industry Engagement”

Direct contacts

Dr. Regina Jutz
University of Mannheim
regina.jutz@uni-mannheim.de

Andreas Hamann
University of Mannheim
andreas.hamann@uni-mannheim.de
Thank you!