Institutional Data Science Services at FDZ UB Mannheim: Enhancing Research Data Management

Renat Shigapov, Irene Schumm, Thomas Schmidt, Jan Kamlah & Larissa Will
13.03.2025, E-Science-Days 2025

Agenda

1. **Introduction**
2. FDZ UB Mannheim
3. Data Science Services
4. Challenges
5. Impact on RDM
6. Conclusions

Introduction: Definition of data science services

- **Data science as a service:** E.g., data collection and modelling as a service
- **Consultancy** on AI and other data science methods
- **Education and training** on reproducible, legal, and ethical use of data science methods
- **IT infrastructure**: GPU servers
- **Community engagement and networking**

Introduction: Existing data science services

All

We

- **Data science services** rapidly emerge at universities worldwide

- **International examples:** Harvard, Utah, Purdue, NC State, and Groningen

- **National example:** Data Science Center at the University of Bremen
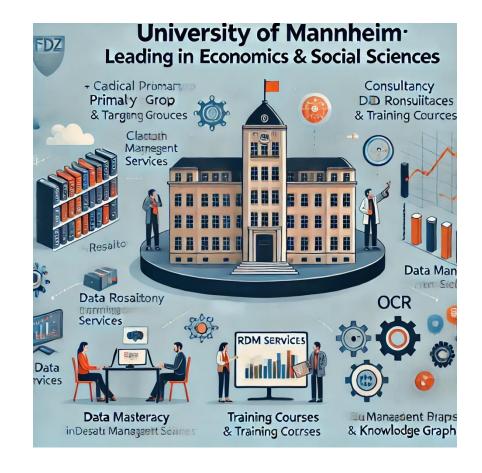
- **NFDI:** NFDI4DataScience

Let's try it too

Agenda

1. Introduction
2. **FDZ UB Mannheim**
3. Data Science Services
4. Challenges
5. Impact on RDM
6. Conclusions

FDZ UB Mannheim: A research data center at Mannheim University Library

- The University of Mannheim is a leading university in economics and social sciences, which is also the primary target group of FDZ

- FDZ offers core RDM services such as a data repository, RDM consultancy and courses, DMP services, data sharing in a secure environment, etc.

- FDZ is involved in seven projects on data literacy & training, RDM, AI, chatbots, OCR, and knowledge graphs

https://fdz.bib.uni-mannheim.de

FDZ UB Mannheim: How we created data science services

FDZ's involvement in the projects on OCR, RDM, and AI enabled us to:

- attract competent staff with AI and RDM expertise and experience
- establish and maintain OCR, RDM, and AI consultancies
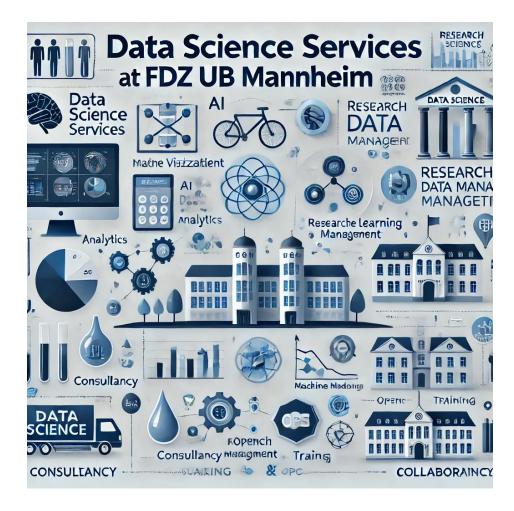- increase AI and RDM competence within the team

https://www.bib.uni-mannheim.de/lehren-und-forschen/forschungsdatenzentrum/fdz-services/data-science-services

Agenda

1. Introduction
2. FDZ UB Mannheim
3. **Data Science Services**
4. Challenges
5. Impact on RDM
6. Conclusions

Data Science Services at FDZ UB Mannheim

Our services include:

1. Data Science as a Service
2. Education & Training
3. Consultancy
4. Community Engagement & Networking
5. Data science tasks at UB
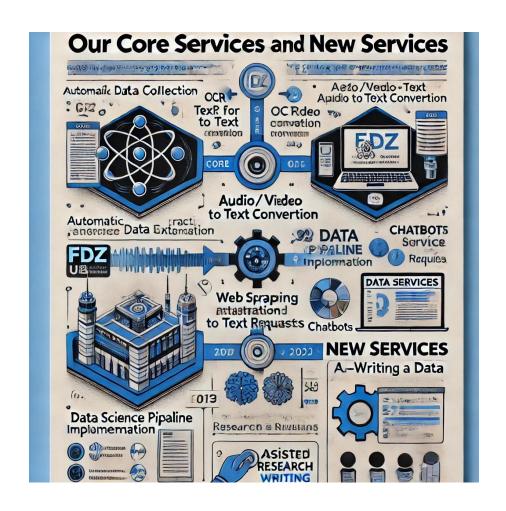
Data Science Services: Data science as a service

Our core services are automatic data collection services:

- OCR

- Audio/Video to Text

- Web scraping and API requests

- Data structuring

New services:

- Data science pipeline implementation

- Chatbots

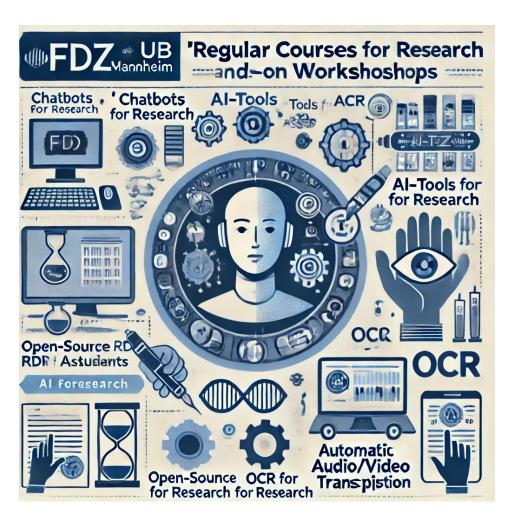- Co-writing a data paper

Data Science Services: Education & Training

Regular courses and hands-on workshops:

- Chatbots for research
- AI-tools for studies and research
- Open-source RDM AI assistants
- OCR for research
- Automatic audio/video transcription

within well-established series:

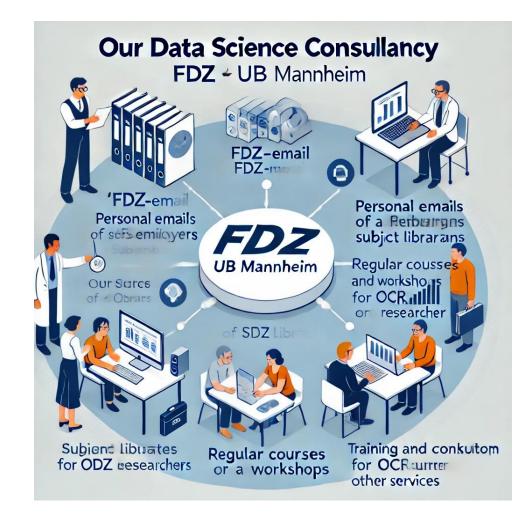- Research skills
- Data literacy essentials
- RDM-seminars

Data Science Services: Consultancy

How do we get customers? Via:

- FDZ-email
- personal emails of FDZ employees
- subject librarians
- courses and workshops
- consultancy hours (e.g., for OCR)
- networking through projects and university events

Data Science Services: Community Engagement & Networking

Uni Mannheim
Academic Speed Dating
on Data Science

organized every
semester by Mannheim
Center for Data Science

GESS (Graduate School
of Economic and Social
Sciences) Research Day

organized by GESS

Data Science Services: Data Science tasks at UB

Chatbots for UB:
- Chatbot about UB
- Chatbots for RDM and open science

Automatic subject indexing:
- DDC
- RVK

Automation of RDM-tasks in FDZ:
- Metadata review
- Creating data documentation

Representation of UB on strategic data science activities of University

Agenda

1. Introduction
2. FDZ UB Mannheim
3. Data Science Services
4. **Challenges**
5. Impact on RDM
6. Conclusions

Challenges

- **Lack of AI strategy:** We built our data science services without the AI strategy of university

  ➡ Now we are involved in creating the AI strategy of Uni Mannheim as well

- **High infrastructural costs:** We need more GPUs to build AI-services, which are so wanted by researchers

  ➡ The University IT plans to buy more GPUs

- **Personnel costs:** Running and maintaining data science services requires skilled staff, making long-term sustainability expensive

  ➡ We need more AI, data science, and RDM projects & internal funding

- **Scalability costs:** "data science as a service" can be very expensive to scale, depending on the specific task

  ➡ Multi-tier support (we might need to become project partners)

Agenda

1. Introduction
2. FDZ UB Mannheim
3. Data Science Services
4. Challenges
5. **Impact on RDM**
6. Conclusions

Impact on RDM

- We enhance data collection at University => More data in our data repository

- Our AI-workshops have more participants than our RDM-workshops => We include RDM-topics in our AI-workshops as well => More outreach

- Automation of our RDM-workflow => Users of data repository get faster reviews for their data/metadata

- Created RDM & Open Science Chatbots => a new service of UB (see the lightning talk by Thomas Schmidt)

Agenda

1. Introduction
2. FDZ UB Mannheim
3. Data Science Services
4. Challenges
5. Impact on RDM
6. **Conclusions**

Conclusions

Data science services are needed at the University of Mannheim:

✔ Our researchers want to outsource at least data collection

✔ UB wants to improve its services with AI

Scalability & sustainability remain key challenges:

✔ Need for GPUs

✔ Long-term funding and skilled staff are crucial

✔ Multi-tier support can help

Community engagement & outreach are highly important:

✔ More data science community events at university are needed

✔ RDM-integrated AI workshops increase researcher engagement

https://fdz.bib.uni-mannheim.de

Help us to improve the Data Science Services at University of Mannheim!

💡 Have ideas to improve our services?

🤝 Interested in collaboration?

📩 Get in touch

Renat Shigapov: renat.shigapov@uni-mannheim.de	Thomas Schmidt: thomas.schmidt@uni-mannheim.de
Irene Schumm: irene.schumm@uni-mannheim.de	Jan Kamlah: jan.kamlah@uni-mannheim.de
Larissa Will: larissa.will@uni-mannheim.de	