AGI 1: Data Literacy in the Physics Curriculum Thursday, March 23, 2023, 11:00–12:30, ZEU/0148

News from PUNCH4NFDI

Dr. Carsten Burgard **PUNCH4NFDI** Education Manager TU Dortmund



technische universität dortmund



What is PUNCH4NFDI?

- PUNCH: Particles, Universe, NuClei & Hadrons
- NFDI: Nationale
 ForschungsDatenInfrastruktur
- Joint consortium of particle, astro-, astroparticle, hadron and nuclear physics
- representing about 9.000 scientists with a Ph.D. in Germany
 - universities
 - the Max Planck society
 - the Leibniz Association
 - the Helmholtz Association.





DAta from PHoton and Neutron Experiments



FAIR data for MATerial Scie

Note: 2 other physics-related consortia



- TA 1: Management and governance
- TA 2: Data management
- TA 3: Data transformations
- TA 4: Data portal
- TA 5: Data irreversibility
- TA 6: Synergies and services
- TA 7: Education, training, outreach, and citizen science

The FAIR principles

- Cornerstone principles of research data management
- achieved to varying degree in the various fields
- unique challenges to be overcome
- culture shift requires long commitments



Skills

- Data Literacy is a keystone skill for modern physicists
- Teaching principles should guide the acquisition of skills
- How is this reflected in physics education?





Status

- Modern university curricula for astronomy and physics must provide skills and methods to cope with large and complex data sets.
- Courses in modern statistics, applied mathematics, or computer science are mostly not mandatory.
- Many curricula lack basic training on software design principles, data structures, sampling methodology, statistics, algorithmic design and optimisation, standards and procedures.
- A systematic training in data science and big data management or modern statistical techniques such as machine learning does not yet exist even on the master or doctoral level.



Goals

- Provide and improve proficiency in NFDI-related themes to enhance career prospects.
- Provide basic educational resources for university-level teaching that will also be offered to other consortia.
 - Integration of topics related to research data management into university curricula, preferably in a commonly recognized core curriculum that will
 - also profit other NFDI initiatives.
- Promote technological literacy matched with a good skill set in data analysis and the adequate understanding of experimental setups.



Work programme

- Conduct a survey among universities to obtain an overview
 - over the conceptual integration of NFDI-related topics into the curricula
 - on the teaching methods and materials,
 - on the extra-curricula activities in the NFDI realm
- Provide recommendations for
 - an NFDI schedule tailored to the needs of bachelor, master and doctoral students.
 - the development and organization of learning materials and tools.
- Offer educational events in the realm of the NFDI-topics.



Work programme

- Enable the use of computing infrastructure for educational purposes.
- Coordinate and administer the education efforts, collect material, propose and discuss topics for schools and workshops, and organize educational events.
- Develop course content and material to be provided to all universities and institutions.
- Coordinate with other community-wide educational programs and resources.



Work packages: Education, training, outreach, citizen science

- WP 7.1: Training of scientists Young Academy
- WP 7.2: Education of students
- WP 7.3: Public outreach
- WP 7.4: Support for citizen science

WP 7.2 Deliverables

- 1. Market survey of available teaching concepts and material.
- 2. Development of standardised curriculum.
- 3. Compilation and development of teaching material for courses and independent learning.
- 4. Aggregation of data resources and access to computing infrastructure.
- 5. Coordination and initiation of educational events, e.g. visiting seminars.



Market survey



- Market survey ("Bestandsaufnahme") being conducted electronically
- not limited to teaching, also includes outreach (categorization sometimes fluid)
- using the DPG LimeSurvey system
 - reliance on DPG resources highly valuable!

Bestandsaufnahme Datenkompetenzen

Part I: Materialien & wiederkehrende Angebote



Part II: Organisationen & Strukturen



Part I: Materialien & wiederkehrende Angebote

- Expect 1 entry per material or course
- 14 features per entry
 - purely objective criteria, this is a data collection!
 - no personal information collected
 - will cover some of the features on the next slide
- Expected participants include
 - organizers & creators
 - teachers & trainers
 - participants & users

Goal: create a catalogue



Part I: Materialien & wiederkehrende Angebote

- Material or course?
- Title
- Organization type (University, School, ...)
- Main topics covered (checkboxes)
- Side topics covered (checkboxes)
- Type / Format (checkboxes)
- target audience (checkboxes)
- intended time to complete
- license
- weblink
- contact person (e.g. organizer)

$\mathbf{\Psi}$	

Part I (contd.). Progress

- Progress of data collection relatively slow
 - only few complete entries so far
 - will leave open & continue to promote
- Only 10-20 complete entries so far
 - you can make a difference!



Please participate! Spread the word!



Part I (contd.). Usage

- Data collected will be made available publicly
- exact mode of hosting not decided yet, but many good options exist
 - e.g. DALIA







Part I (contd.). Future plans

- Sanitize data
 - remove duplicates, fill missing info
 - o manual work, but hopefully manageable
- Compose catalogue
 - host on DALIA or similar systems
- Plan to follow up with evaluation survey
 - How well did participants receive the material?
 - How widely is it used?

Goal: Establish best practices!



Part II: Organisationen & Strukturen

- Expect 1 entry per organization
- 11 features per entry
 - purely objective criteria, this is a data collection!
 - no personal information collected
 - will cover some of the features on the next slide
- Expected one responsible person from the organization to participate

Goal: Identify big players!



Part II: Organisationen & Strukturen

- Name
- Research Institution Identifier
- Type (University, School, ...)
- Main target audience
- Main topic (checkboxes)
- Side topic (checkboxes)
- Types of offers (checkboxes)
- weblink
- contact person

	Φ	

Part II (contd.). Progress

- Progress of data collection even slower
 - as expected, only few organizations
 - promotion even more important
 - \circ less than 5 replies so far



Please participate! Spread the word!



We are not alone!





Workshop April 2022: Research data in physics studies

- Fundamental issue: the importance of data issues is increasing rapidly \rightarrow <u>but</u> no corresponding attention in physics studies
- Idea:
 - Networking between "Deutsche Physikalische Gesellschaft" (DPG), "Konferenz der Fachbereiche Physik" (KFP) and physics-related NFDI consortia
 - Identify problem areas and fields of action
 - Discuss good practise examples
 - Develop ideas/proposals for adapted physics curricula









Summary & Conclusions of the workshop

- Many different big players presented themselves
 - NFDI sections PUNCH, FAIRmat
 - DAPHNE
 - 0 ...
- Many topics covered, from central data hubs to electronic lab reports
- Discussion points
 - definitions of key terms
 - position statements by ZapF, KFP
- Very unanimous conclusions that there is a lot of potential for improvement
 - Key decision: collect best practices!

Live protocol of the sessions:



Idea: Create a whitebook!

- common statement across NFDI consortia (section edutrain), DPG, KFP, ZapF
- Make recommendations
- State importance of the topic
- establish best practices
- defining targets
- discuss different scenarios
- compare across different courses: Bachelors of Science / Education, Master of Science / Education
- collect literature



Phone interviews

- Conduct phone interviews to learn about teaching activities at universities
- interview covers 8 topics, around 40 minutes in total
 - programming languages
 - statistics
 - machine learning
 - FAIR principles
 - electronic lab books
- progress is slow but steady
 - hope for completion by summer

how do we prioritize data literacy versus other keystone skills in the physics curriculum?



Conclusions

- PUNCH4NFDI is very active in various initiatives
 - data collection is ongoing, too early to draw conclusions
 - wealth of different projects (e.g. DALIA) is ramping up
 - difficult to stay informed
 - a lot of active networking events ongoing NFDI "Kartierungs-Workshop" (Feb)
- Please participate in the surveys & spread the word!





Bestandsaufnahme Datenkompetenzen



Thank you!

Workshop: Lay of the Land

- NFDI workshop
 <u>nfdi.de/datenkompetenz-workshop/</u>
- Lay of the land ("Kartierung")
- many different types of projects
- several of them dedicated to hosting & collecting teaching materials

Miro Board with summary of particpants



