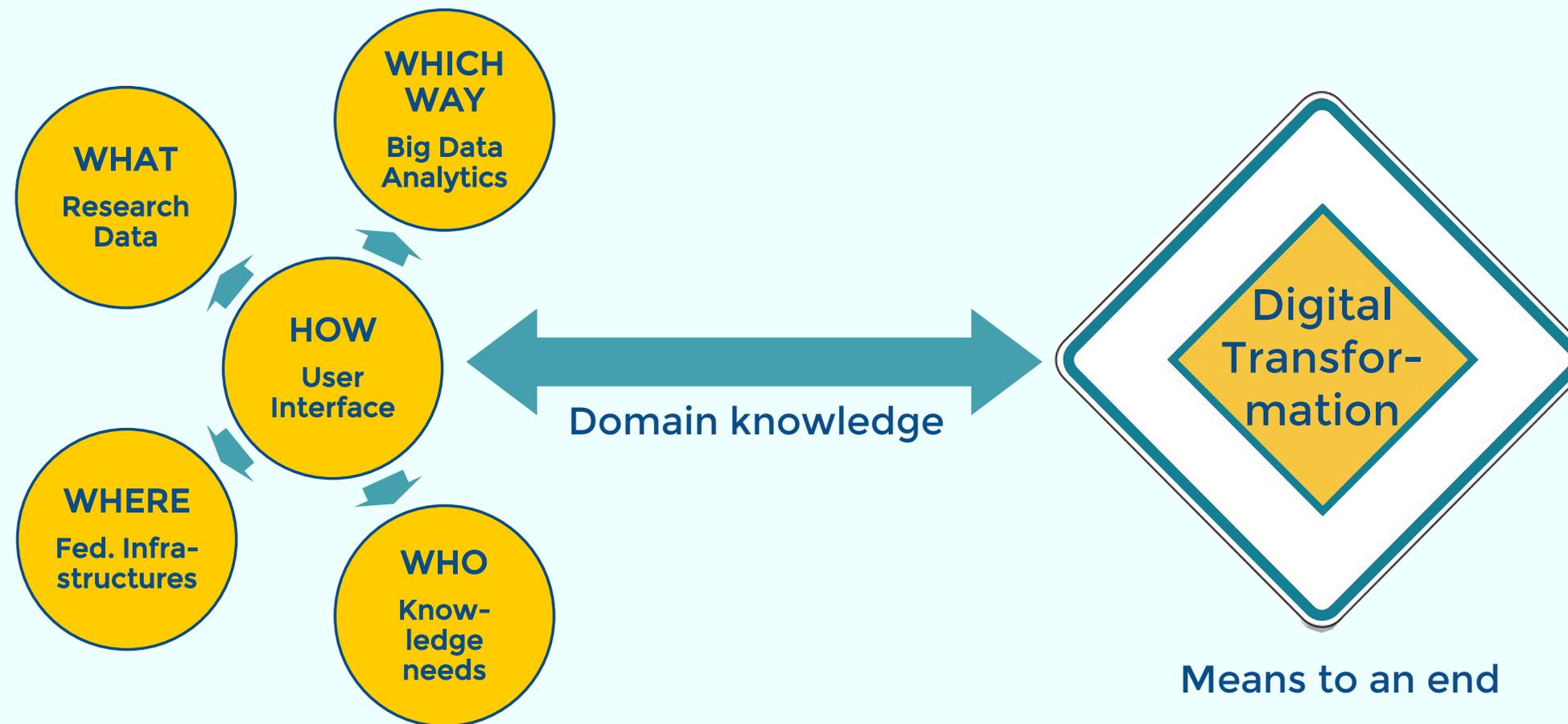
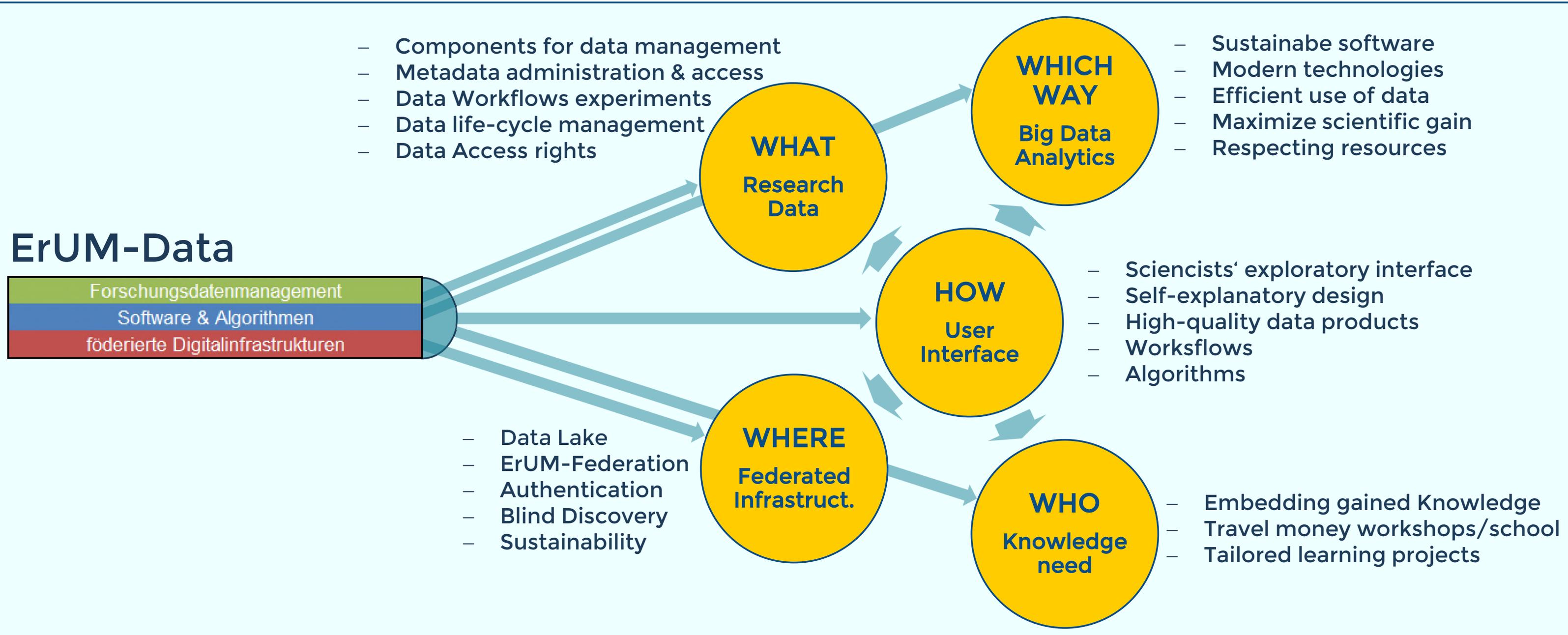


DIG-UM & ErUM-Data-Hub

Objective: Accelerate Scientific Discoveries in Research on Universe & Matter



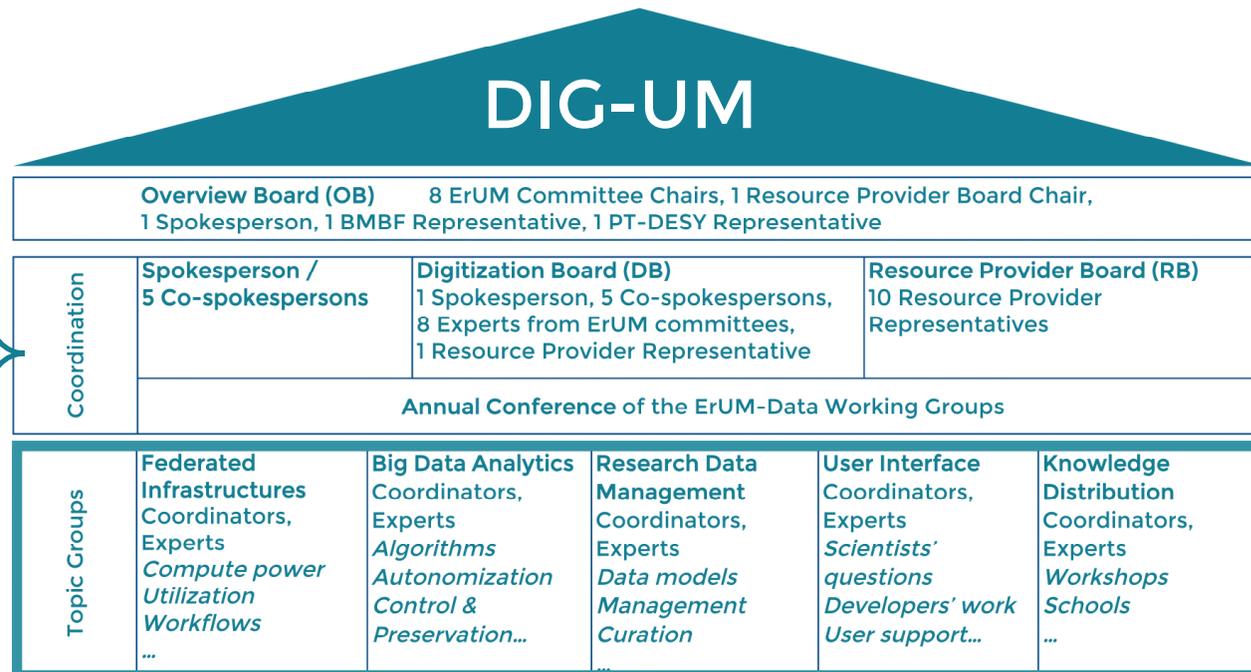
Connection to ErUM-Data Themes



Structure: Digital Transformation

Community Self-Organization

KAT
KET
KfB
KFN
KFS
KFSI
KHuK
RDS



BMBF Funding

ErUM-Pro
Project funding for Experiments

ErUM-Data
120 M€ / 10 years
Software & Algorithms
17,5M€ / 3 years

...



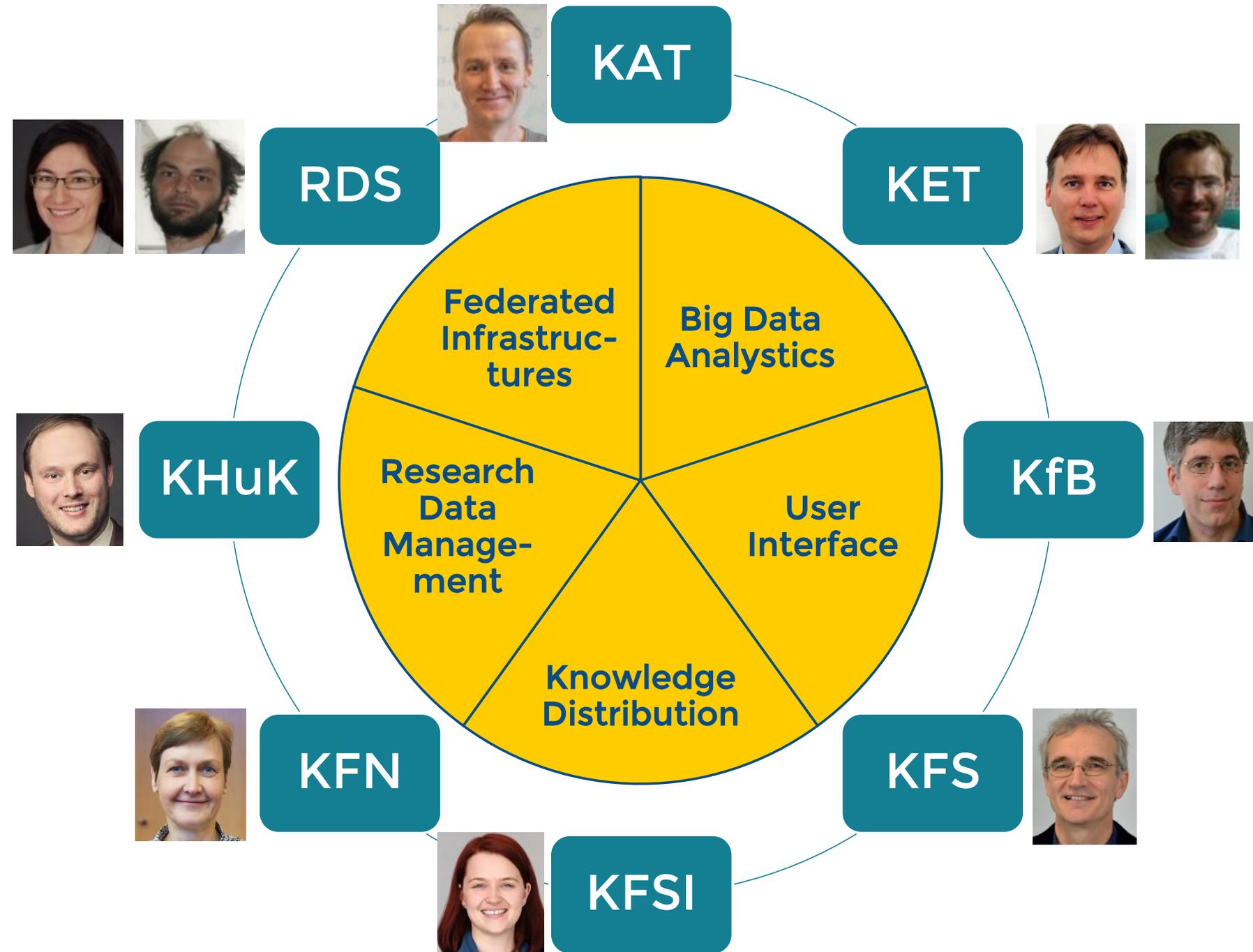
ErUM-Data-Hub
2,5M€ / 4 years

Komitee

KAT Astroteilchenphysik
KET Elementarteilchenphysik
KfB Beschleunigerphysik
KFN Forschung mit Neutronen
KFS Forschung mit Synchrotronstrahlung
KFSI Forschung mit nuklearen Sonden und Ionenstrahlen
KHuK Hadronen- und Kernphysik
RDS Rat Deutscher Sternwarten

Topic Group Chairs & Deputies

with Digitization Experts
from all ErUM Communities

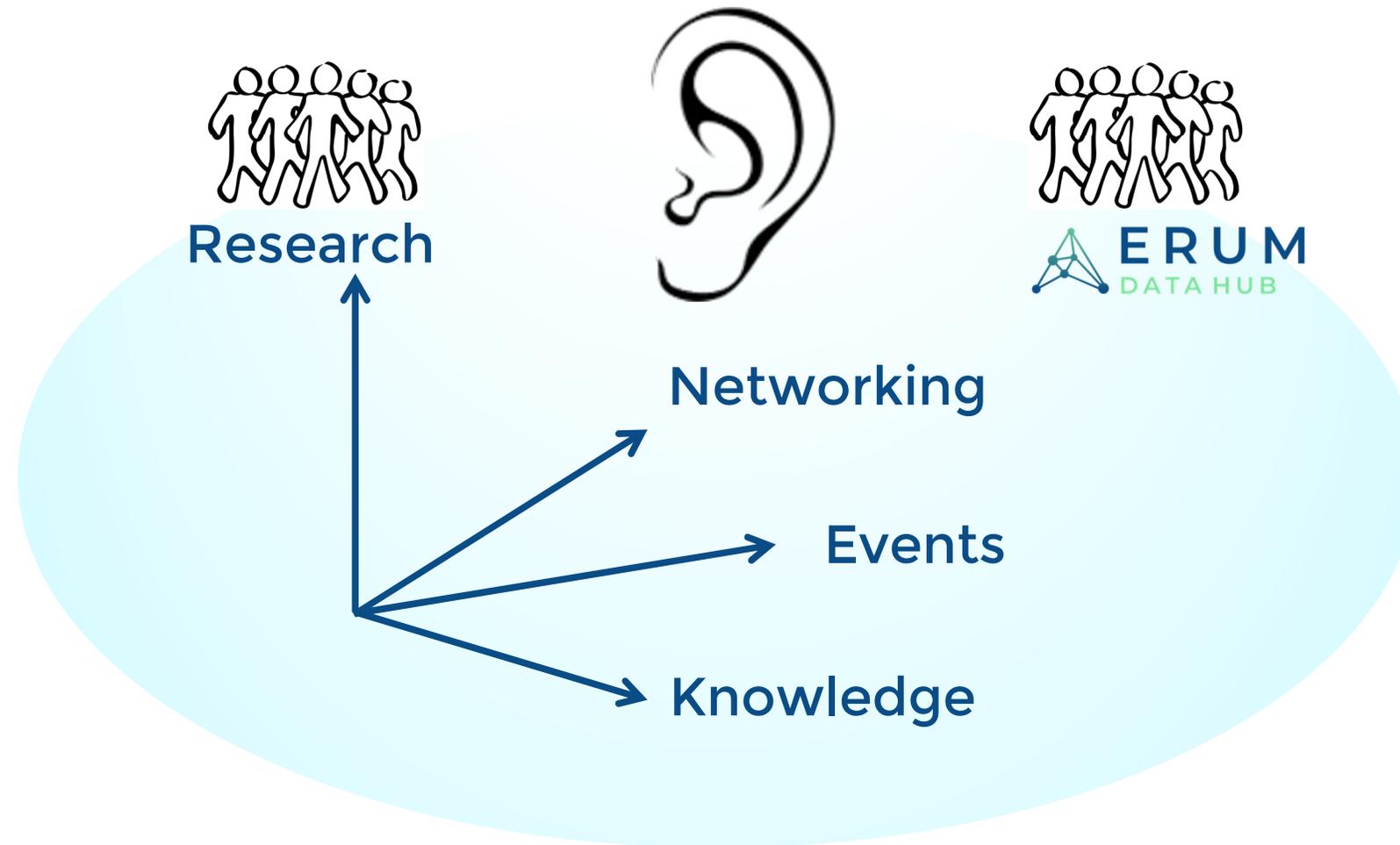


- Regular Meetings
- Strategies
- Recommendations
- Workshops & Programmes

Tight Relation: DIG-UM ↔ ErUM-Data-Hub



KAT
KET
KfB
KFN
KFS
KFSI
KHuK
RDS



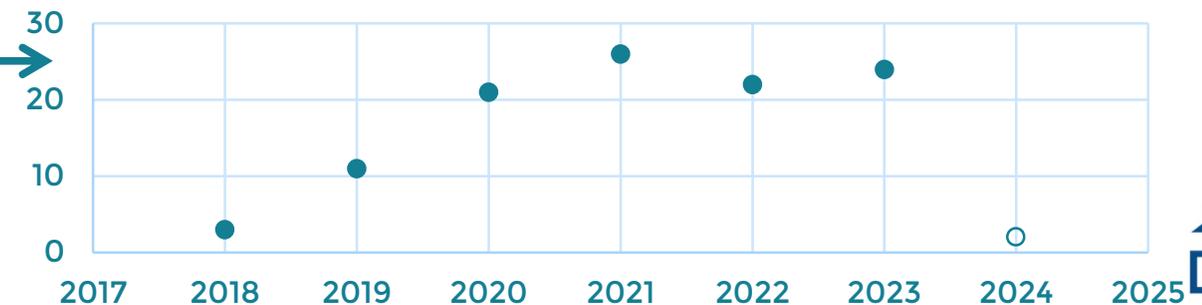
In addition regular exchange

- Hub-Team coordinations
- DAPHNE4NFDI
- PUNCH4NFDI
- Terascale, Astronomy-Schools, Teilchenwelt,...
- LHC-Belle-Offices
- RWTH AI-Center
- ...

DIG-UM

Overview Board (OB) 8 ErUM Committee Chairs, 1 Resource Provider Board Chair, 1 Spokesperson, 1 BMBF Representative, 1 PT-DESY Representative					
Coordination	Spokesperson / 5 Co-spokespersons	Digitization Board (DB) 1 Spokesperson, 5 Co-spokespersons, 8 Experts from ErUM committees, 1 Resource Provider Representative	Resource Provider Board (RB) 10 Resource Provider Representatives		
	Annual Conference of the ErUM-Data Working Groups				
Topic Groups	Federated Infrastructures Coordinators, Experts Compute power Utilization Workflows ...	Big Data Analytics Coordinators, Experts Algorithms Automatization Control & Preservation ...	Research Data Coordinators, Experts Data models Management Curation ...	User Interface Coordinators, Experts Scientists' questions Developers' work User support ...	Knowledge Distribution Coordinators, Experts Workshops Schools ...

DIG-UM Organizational Board Meetings



Networking - Transfer Office serving Digital Transformation in Research on Universe & Matter



Knowledge repositories

www.erumdatahub.de

wiki.erumdatahub.de



Networking

Enabling
Knowledge

Expert Workshops
Schools

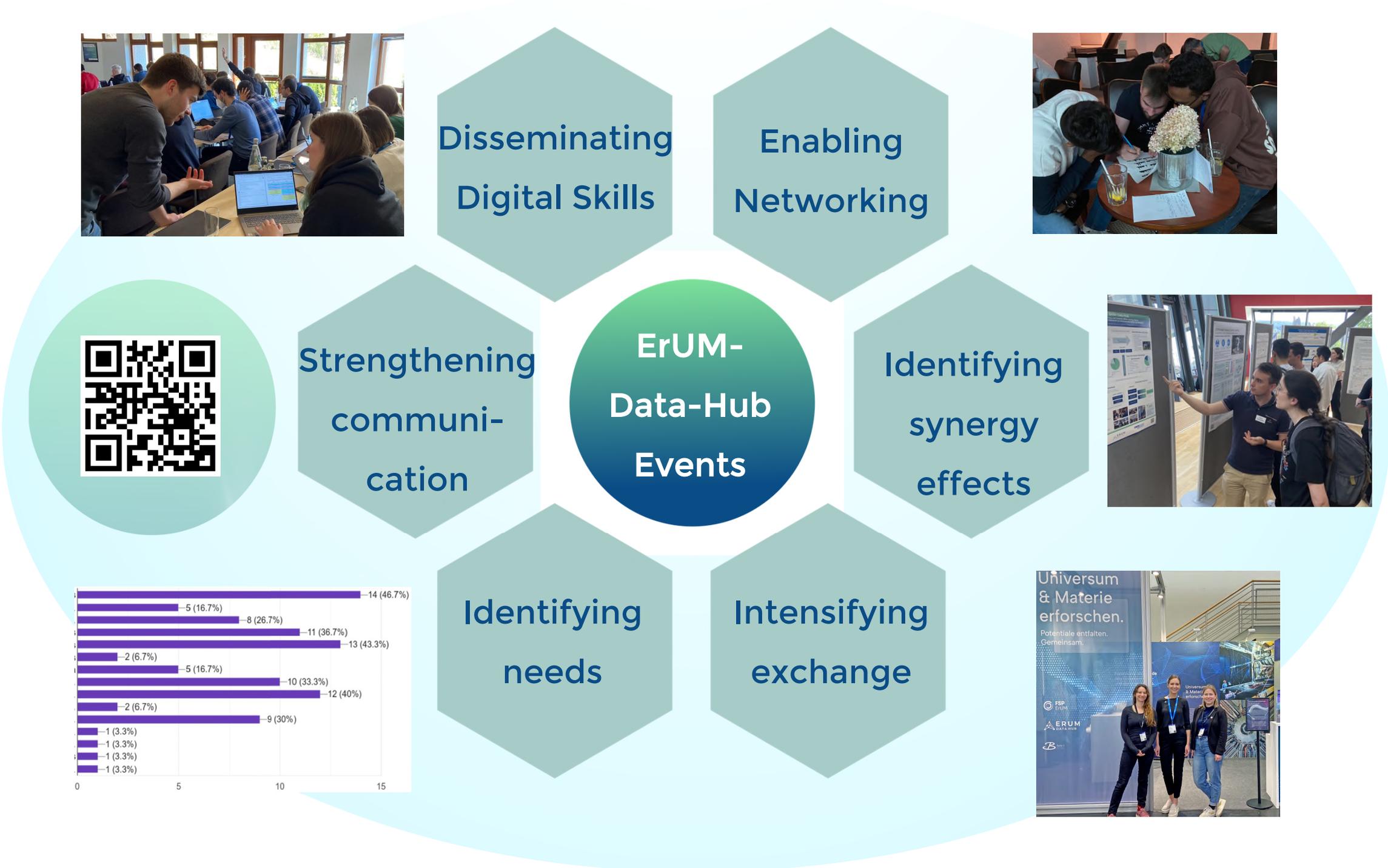


Transfer

Hannover
Messe

Full-time: Angela Warkentin (Team Leader), Judith Steinfeld (Communication)
Part-time: Peter Fackeldey, Benjamin Fischer, Stefan Fröse, Ulla Lardinoix, M.E.

ErUM-Data-Hub Events: Holistic approach on Knowledge Transfer & Networking & Synergies



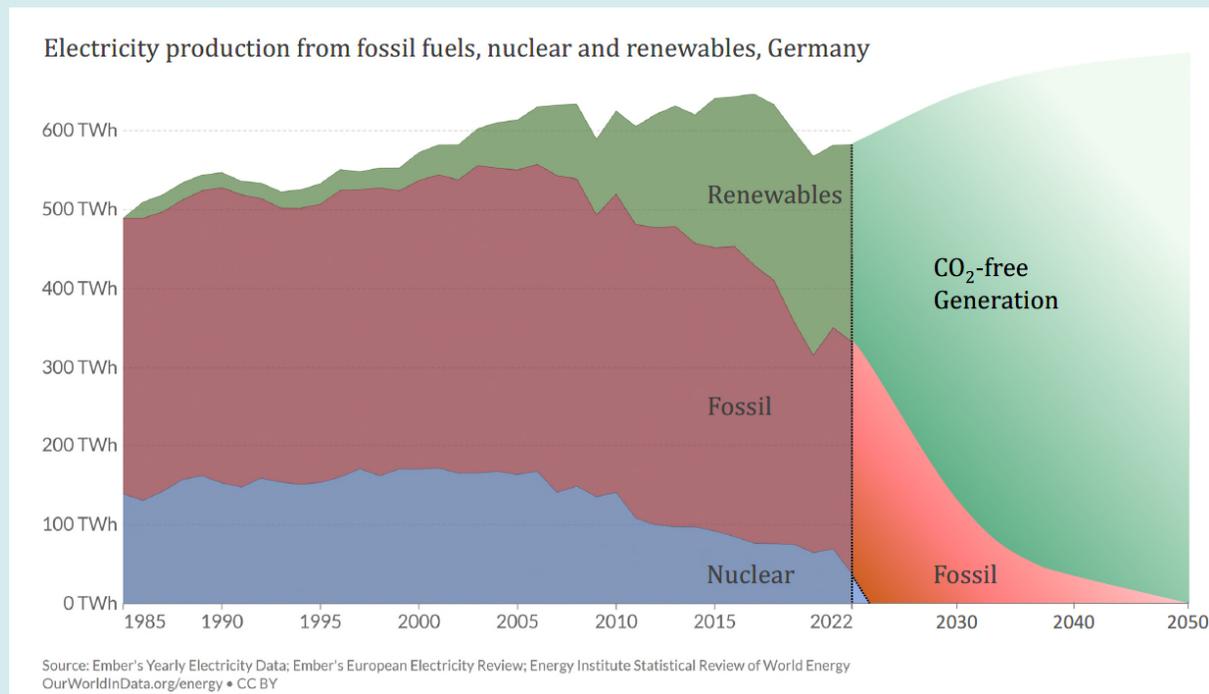
Sustainability in ErUM Research

May 23 Workshop : Awareness sustainable science

Knowledge gain ↔ Resource usage

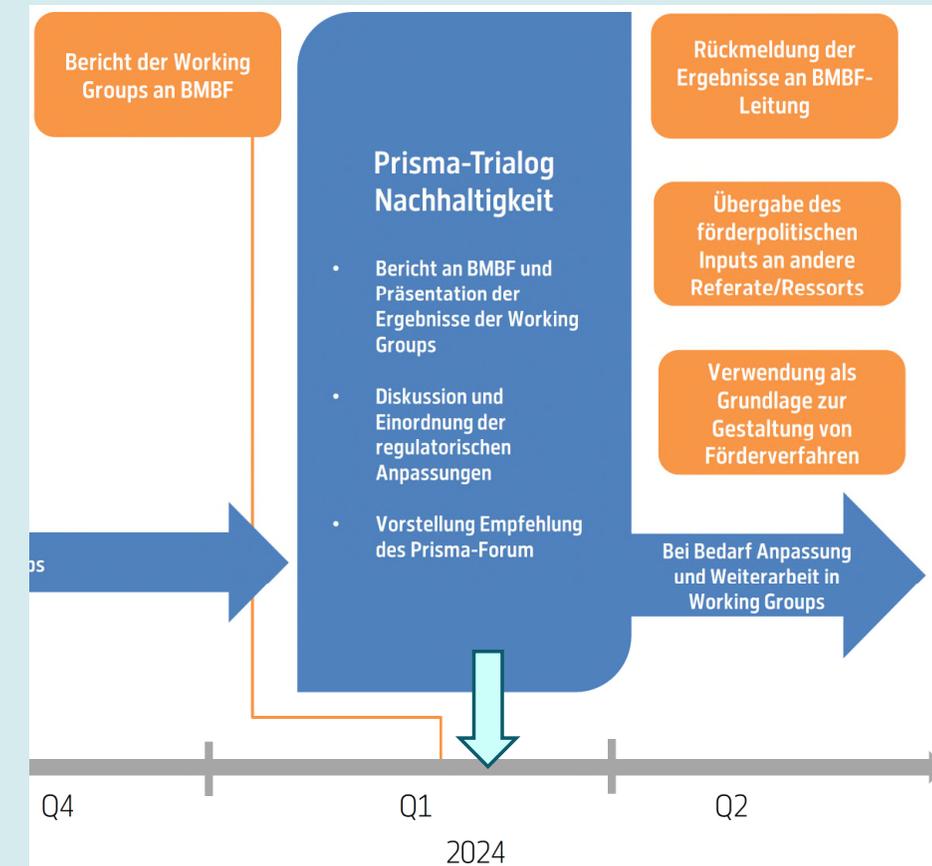
<https://arxiv.org/abs/2311.01169>

Portfolio of measures: short-medium-long term



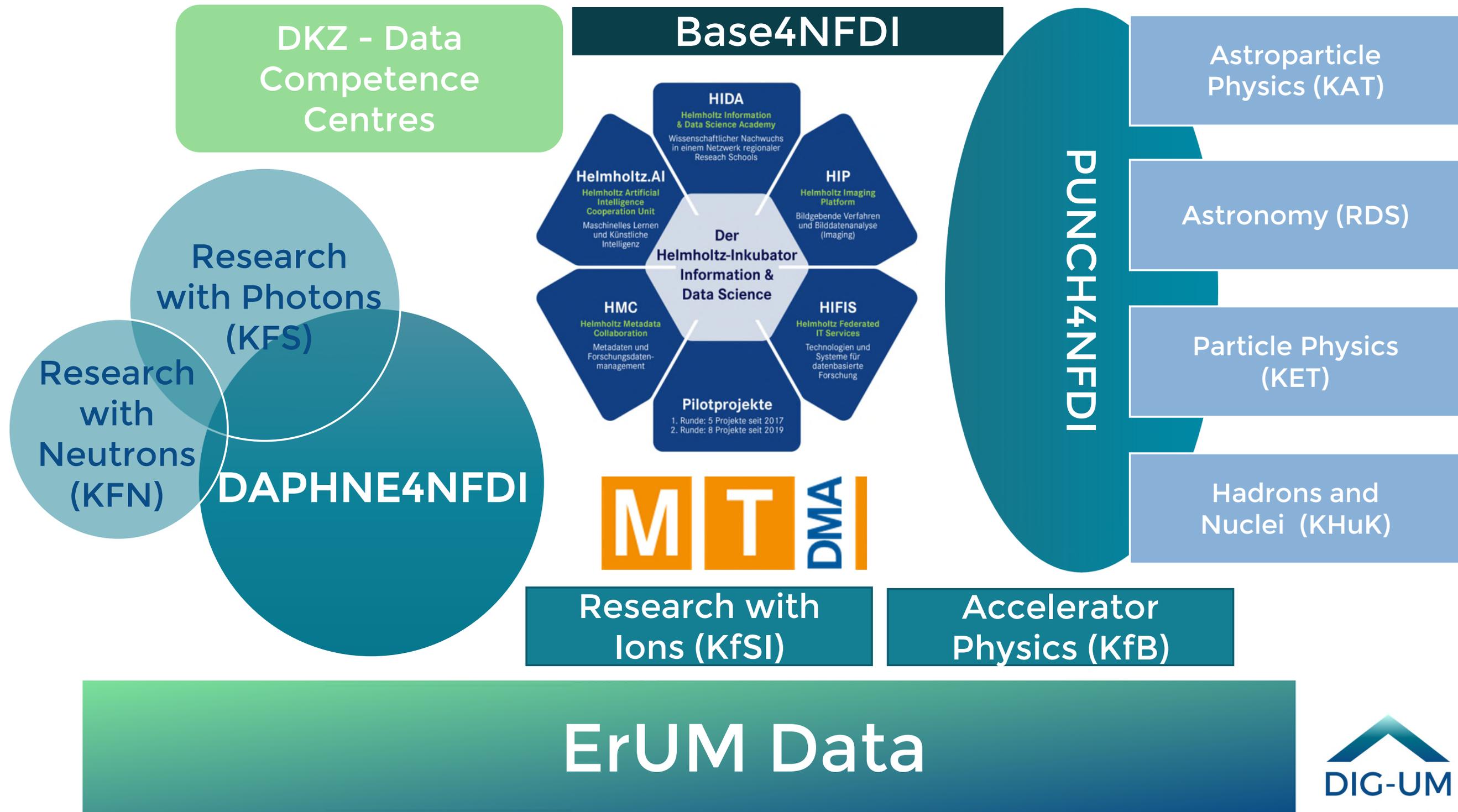
Since May 23 BMBF PRISMA Trialog „Nachhaltigkeit in der Forschung an Großgeräten: Ressourceneffizienz & Zukunftssicherung“

S. Bohleber



Multiple Connections on Digital Transformation

International



Multiple Connections on Digital Transformation

ErUM-Data interacts with multiple connections:

- ErUM-Data acting in view of eight heterogeneous Communities, developing solutions to multiple challenges arising in ErUM research fields.
- A lot of communication, especially with NFDI to diminish gray zones: Profit from work done by others, avoid duplication or contradictions.
- ErUM-Data / NFDI closely follow international developments.

→ Multiple funding programs are complementary and urgently needed.

DIG-UM question to ErUM-Data Call (1)

How they interplay:

ErUM-Pro & ErUM-Data

- **Will there be funding for experiment-related developments that are essential for the success of the experiment?**

DIG-UM question (2)

Community Stages

Digitization needs and development status vary strongly between communities.

➤ Will the evaluation take this into account?

DIG-UM question (3)



Application Fields are Research Data Management, Software & Algorithms, Federated Digital Infrastructures:

Developments within Topic Group User Interface concern all three fields.

➤ **Can all three fields be integrated in a single application?**

DIG-UM question (4)

Enabling sustainability of entire structure: computing resources, maintenance & development of software data management solutions, etc. Situation challenged by 3-years project funding.

- **Do you see a new gateway opening to finance experienced long-term employees?**
- **Could a “Transition-Funding” for knowledge transfer to next-generation employees be considered? Under which circumstances could this be funded?**

DIG-UM question (5)

ErUM-Transfer

Within ErUM: Transfer between ErUM communities.

Outside ErUM: Opportunities for new developments with multiple applications outside ErUM research.

➤ **Will - or should - transfer play a role in the next ErUM-Data funding cycle?**

Formalities of upcoming ErUM-Data call (6)

- **Application stages:** one-stage application process?
- **Application language:** project description in English, AZAPs in German?
- **Consortium Size:** Lessons from previous funding cycle? E.g. cross-community applications require substantial size.
- **Cooperation:** competitive groups versus cooperative complementary groups?
- **Theory:** Can theory groups be included?

Members within DIG-UM rather don't like general requirements for applications such as:

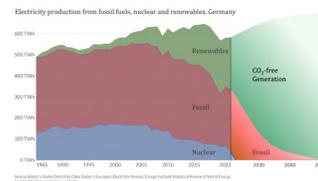
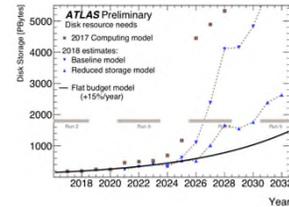
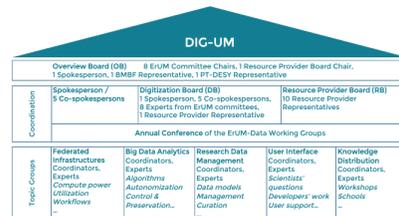
- **Matter, Universe, Particle:** Applications from two research fields
- **Non-university Partners:** Partner from Industry
Non-university Partners: Partner from large scale research infrastructure
- **Computer Scientists:** most welcome, however, their research interest rarely matches application development within ErUM research.

DIG-UM question (7)

Expert Panel

- **Could experts from the field of digitization who have substantial domain knowledge of ErUM research be included?**

Summary [1/4]



- **DIG-UM Community: well alive & very active.**
- **Major challenges ahead:**
 - We must prepare for a future with exponentially increasing data volumes.
 - Exploiting the upcoming science data constitutes major compute challenges to us (hardware, services, algorithms).
 - Our digital transformation is currently too slow compared to US and others - with impact far beyond our research.
 - Sustainability measures constitute major challenges to ErUM.

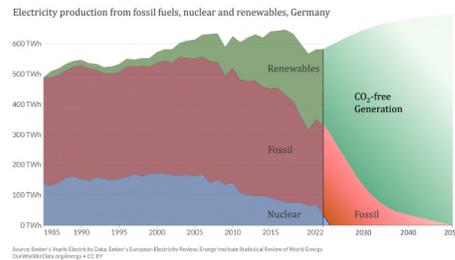
Summary [2/4]: ErUM-Data-Hub



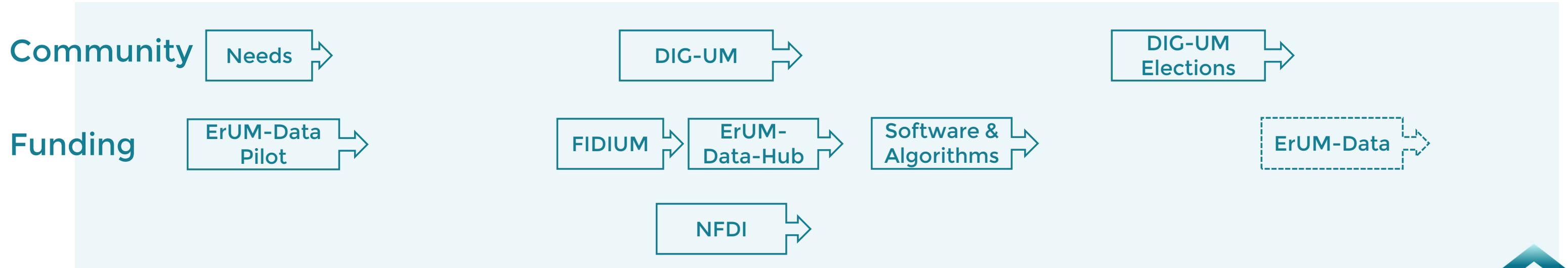
ErUM-Data-Hub: central to digital transformation in ErUM.

Recommend Phase 2 of ErUM-Data-Hub:

- 1. Accelerate cross-community solutions for**
 - a. Data and computing challenges, and**
 - b. Sustainability measures.**
- 2. Distributed concept to enlarge reach and impact.**



Summary [3/4]: Time Line Digital Transformation



Wir danken dem BMBF für das herausragende Engagement für die ErUM-Wissenschaften, für die Initiative, die digitale Transformation in der ErUM-Forschung über den Aktionsplan ErUM-Data voranzubringen, und für die substantielle finanzielle Unterstützung.

Unser Dank gilt auch dem Projektträger DESY für kontinuierliche und konstruktive Unterstützung.

Backup

ErUM-Data-Hub Events: Holistic approach on Knowledge Transfer & Networking & Synergies



Providing training in the latest technologies for ErUM-scientists

Disseminating Digital Skills

Enabling Networking

Event program with focus on networking aspect



Target group-oriented documentation of events and results

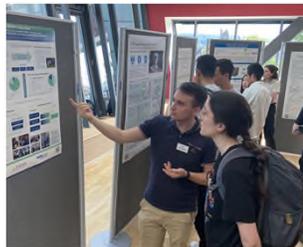


Strengthening communication

ErUM-Data-Hub Events

Identifying synergy effects

Scientific advisors (Referent:in) tasked with expanding the network and establishing contact between ErUM communities & projects

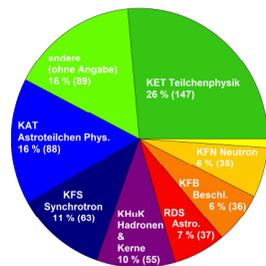


Needs assessment in the form of surveys and discussions to evaluate and adjust medium to long-term orientation

Identifying needs

Intensifying exchange

Related scientific fields & industry (innovation and knowledge transfer)



Phase 2: ErUM-Data-Hub

Sustainability

- Service Center advances ErUM-Data sustainable research.

Scientific Advisors (Referent:in)

- Outstanding achievements by part-time PhD students.
- Experienced-and-qualified postdocs (?)

Distributed ErUM-Data-Hub

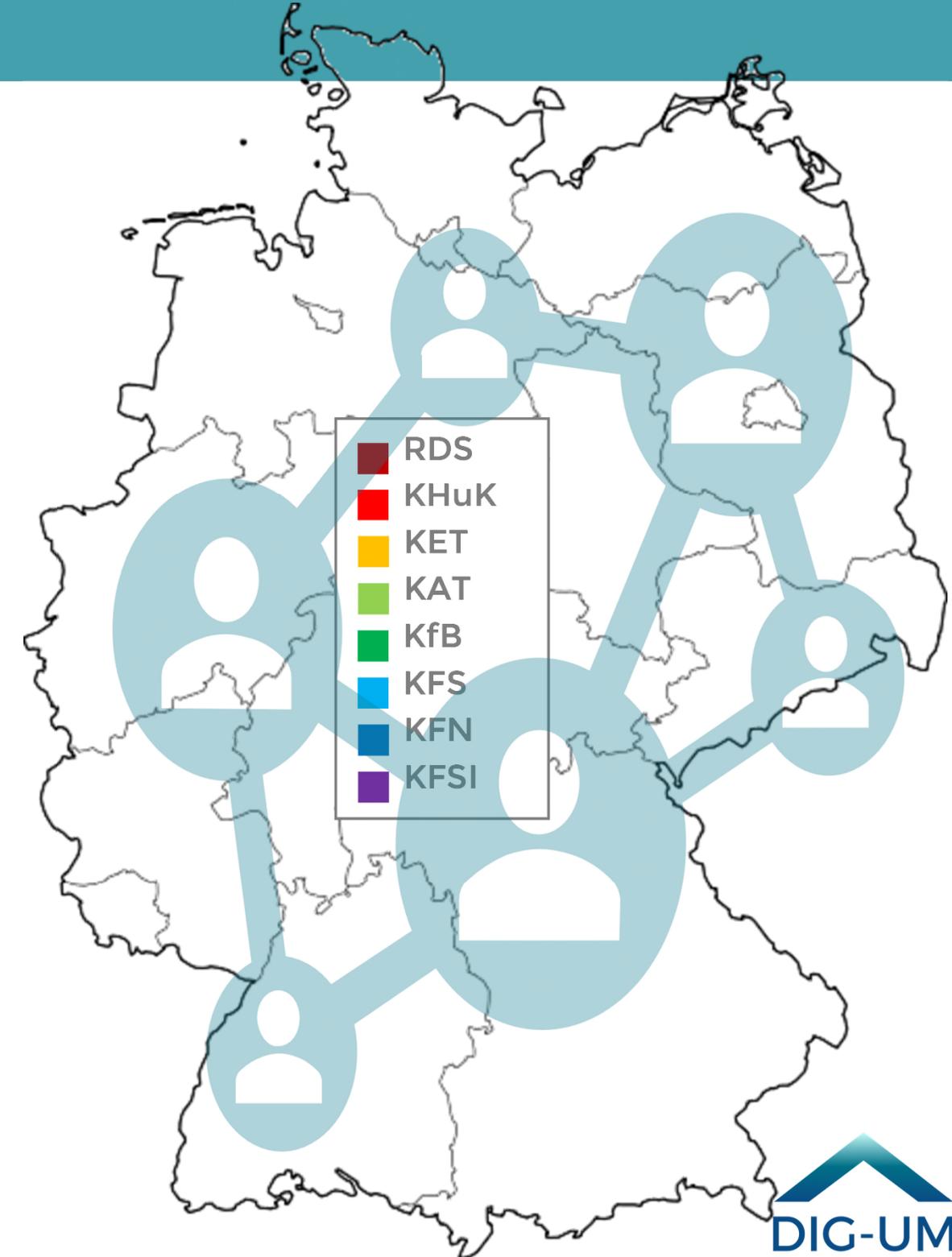
- Bringing communication & workshops to people.
- Involve ErUM PIs/institutes all over GERMANY.

Reach 20,000 Scientists

- Current reach through ErUM committees & institute leaders.
- Innovative direct communication with young scientists.

Support ErUM-Data-Hub PI / Spokesperson

- Lectureship position substituting lectures, supervising scientific group.



Overview Board 12-Dec-2023

Committee		Representative	Deputy
Forschung mit Synchrotronstrahlung	KFS	Christian Gutt	Bridget Murphy
Rat Deutscher Sternwarten	RDS	Stefanie Walch-Gassner	Volker Springel
Hadronen- und Kernphysik	KHuK	Michael Block	Anton Andronic
Elementarteilchenphysik	KET	Markus Schumacher	Lutz Feld
Forschung mit Neutronen	KFN	Mirijam Zobel	Sabrina Disch
Astroteilchenphysik	KAT	Uli Katz (Vorsitz)	Kathrin Valerius
Beschleunigerphysik	KfB	Erik Bründermann (Deputy)	Florian Hug
Forschung mit nuklearen Sonden & Ionenstrahlen	KFSI	Günther Dollinger	Daniel Severin
Guests:			
BMBF	BMBF	Simon Bohleber	
DESY Projektträger	PT.DESY	Salome Shokri-Kuehni	Sarah Bühler
Chair Resource Provider Board		Günter Duckeck	
Spokesperson		Martin Erdmann	
Scientific Secretary		Dirk Lützenkirchen-Hecht	

Digitization Board from 20-Dec-2023

Committee		Representative	Deputy
Forschung mit Synchrotronstrahlung	KFS	Dirk Lützenkirchen-Hecht	Bridget Murphy
Rat Deutscher Sternwarten	RDS	Marcus Brüggem	Tobias Buck
Hadronen- und Kernphysik	KHuK	Tobias Stockmanns	Sören Lange
Elementarteilchenphysik	KET	Günter Quast	Christian Zeitnitz
Forschung mit Neutronen	KFN	Michael Schulz	Andreas Houben
Astroteilchenphysik	KAT	Andreas Haungs	Gernot Maier
Beschleunigerphysik	KfB	Erik Bründermann	
Forschung mit nuklearen Sonden & Ionenstrahlen	KFSI	Judith Reindl	
Spokesperson		Martin Erdmann (Spokesperson)	
Topic Group Federated Infrastructure		Markus Demleitner*	Kilian Schwarz
Topic Group Big Data Analytics		Thomas Kuhr*	Jan Steinheimer
Topic Group Research Data Management		Astrid Schneidewind*	Monica Valencia-Schneider
Topic Group User Interface		Pierre Schnizer*	Tim Ruhe
Topic Group Knowledge Distribution		Dirk Lützenkirchen-Hecht*	Judith Reindl
Representative Resource Provider Board		Günter Duckeck	Thorsten Kollegger
ErUM-Data-Hub (Guest)		Angela Warkentin	
ErUM-Data-Hub (Guest)		Peter Fackeldey	
ErUM-Data-Hub (Guest)		Benjamin Fischer	

Resource Provider Board 1-Jan-2024

Computing Site	Representative	Deputy
DESY	Volker Gülzow	NN
GSI	Thorsten Kollegger (Chair)	Dmytro Kresan
KIT	Achim Streit	Andreas Petzold
AIP	Harry Enke	Arman Khalatyan
TIER 2 Sites:		
Aachen	Alexander Schmidt	Thomas Kress
Freiburg	Markus Schumacher	Michael Böhler
Göttingen	Arnulf Quadt	Sebastian Wozniewski
LMU München	Günter Duckeck	Otmar Biebel
MPI München	Stefan Kluth	NN
Wuppertal	Christian Zeitnitz	Marisa Sandhoff

ErUM-Data-Hub Projekt: 5 Arbeitsbereiche

Arbeitsbereich 1: Vernetzungsmaßnahmen: ErUM-Communities im Bereich Digitalisierung

AP1: Einrichten des ErUM-Data-Hubs

AP2: ErUM-interne Zusammenarbeit

AP3: Organisatorischer Support zur IT-Infrastrukturentwicklung

Arbeitsbereich 2: Austausch stärken: Technologie und Wissenstransfer

AP1: Zusammenarbeit mit verwandten Wissenschaftsbereichen

AP2: Zusammenarbeit mit Wirtschaft

AP3: Vernetzung mit wissenschaftlichen Organisationen

Arbeitsbereich 3: Digitale Kompetenzen ausbauen: Aus- und Weiterbildung

AP1: Infrastruktur

AP2: Basisausbildung ErUM Data Nachwuchs

AP3: Vertiefende Ausbildung ErUM Data

AP4: Weiterbildung und Austausch zwischen Expert*innen

Arbeitsbereich 4: Kommunikation stärken: Hub-Kommunikation

AP1: Outreach und Öffentlichkeitsarbeit

AP2: Citizen Science

AP3: Mediale Ereignisse zu Beginn und Abschluss der Förderlaufzeit

Arbeitsbereich 5: Unterstützung in Administration und IT

Aufbau und Betrieb des Büros

Aufbau und Betrieb von Interaktionszentren

Pflege und Betreiben erstellter Wissensspeicher

Betrieb der Webplattform für ErUM-Data

Betrieb der online-Plattform für den Austausch in der Aus- und Weiterbildung

Übergabe der erstellten Wissensspeicher an die ErUM-Data Communities

- AP: Arbeitspaket

BMBF: Informeller Austausch ErUM-Data-Hub 8.3.22

Teilnehmende

Prisma-Forum:

Prof. Dr. H. Dosch
Prof. Dr. E. Elsen
Prof. Dr. R.-D. Heuer
Prof. Dr. B. Kanngießer
Prof. Dr. S. M. Schmidt
Prof. Dr. M. Tolan
Prof. Dr. J. Wambsganß

ErUM-Data-Hub:

Prof. Dr. M. Erdmann
Prof. Dr. U. Katz
Dr. B. Murphy

Zuwendungsgeber:

Dr. R. Dieter
Dr. V. Dietz
Dr. W. Ehrenfeld
Dr. D. Hoppe
Dr. D. Konold-von Bornstaedt
E. Lilienthal

Vernetzung der ErUM-Fachgemeinschaften

Führt Aktivitäten der Fachgemeinschaften zusammen

Zusammenarbeit von Hochschulen, Forschungseinrichtungen und
Forschungsinfrastrukturbetreibern (FIS):

Soll verstärkt auf Hochschulbedarfe eingehen, Helmholtz-Plattformen einbeziehen

Hub-Arbeitsbereich Digitale Kompetenzen / Aus- und Weiterbildung

Aktivität soll mit hoher Priorität verfolgt werden

Frühe Ausbildungsstufen sollen einbezogen werden

Zusammenarbeit mit der NFDI

Aktivitäten von NFDI und ErUM-Data müssen eng abgestimmt werden

Neue Hardware-Strukturen

Einsatz in ErUM soll mit Unterstützung des Hubs frühzeitig evaluiert werden

Weitere Empfehlungen

Erwartungsmanagement: erreichbare Ziele für ErUM-Data-Hub ansetzen

Förderierte Infrastrukturen in ErUM-Data

M. Demleitner, K. Schwarz und die
Dig-UM TG Federated Infrastructures

Prisma-Strategiegespräch
ErUM-Data 2025-2028



Föderierte Digitale Infrastrukturen

Aspekte der Förderung

Zusammenfassung und Ausblick

Föderierte Digitale Infrastrukturen

FDI sind verteilte
Rechnersysteme, die
Forschungsdaten

- nach gemeinsamen
Standard-Protokollen
- vernetzt
- bei Bedarf mit einheitlicher
Authentifikation
- und einheitlich entdeckbar

verarbeiten oder zur Verfügung
stellen, um so deren

- Produktion,
- Analyse,
- Speicherung,
- Entdeckung und
- Verbreitung

zu erleichtern oder zu
ermöglichen.

- Einheitliche Schnittstellen und Policies sorgen für einfachen Zugang zu verschiedenen Ressourcen, v.a. bei kurzfristigen Spitzenbedarfen
- Daten-Wiedernutzung (und Open Data) sind frei von technischen Hürden und werden durch Discovery-Services erleichtert
- Kein Lock-in auf bestimmte Werkzeuge: Software baut auf wohldefinierten Schnittstellen auf.

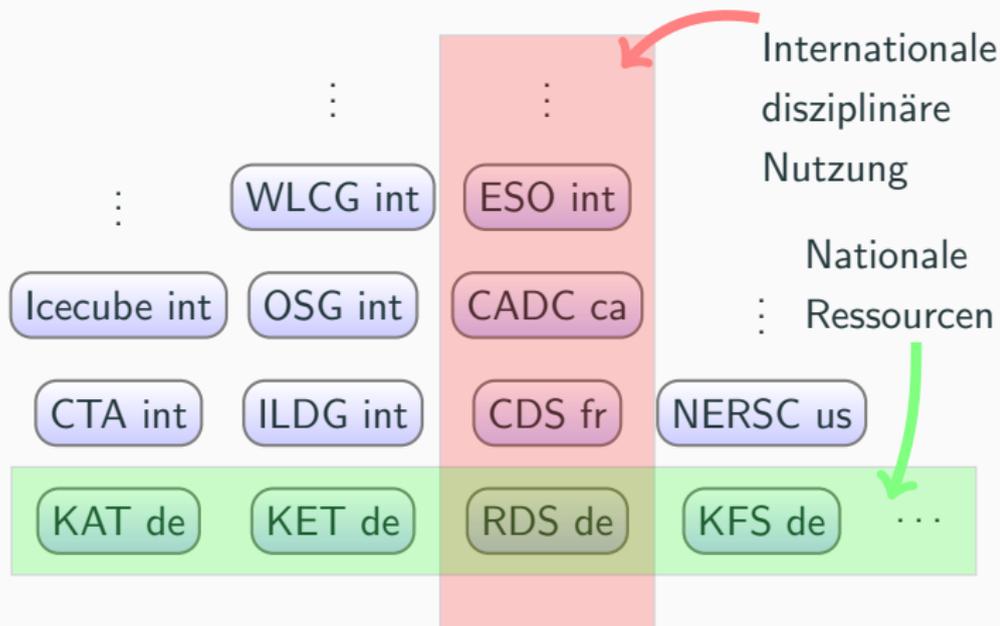
- Gemeinsam genutzte förderierte Ressourcen führen zu effizienterer und nachhaltigerer Auslastung
- Geringerer administrativer und finanzieller Aufwand durch gemeinsamen Betrieb und gemeinsam entwickelte Standardsoftware

Aber...

- Die Definition und Entwicklung von Standards ist aufwändiger als Einzellösungen
- Zwänge zur Bedienung projektspezifischer Interessen über die der Gesamt-Community.

Deshalb: Föderation braucht langfristige und kontinuierliche Förderung.

Die internationale Einbettung



Für Wissenschaftler_innen relevante Föderation muss mit deren internationaler Ausrichtung verträglich sein.

Aspekte der Föderierung

Abgesehen von öffentlichen Archiven braucht praktisch jede Infrastruktur Logins („Authentication“), und es ist für Betreiber wie Nutzende unbequem, wenn diese für jede Ressource separat vergeben werden.

Auch müssen Rechte zum Zugriff auf Ressourcen transparent und einheitlich vergeben werden können („Authorization“).

Fast alle Disziplinen melden Bedarf an förderierter AAI (insbesondere Single Sign On) an.

- X.509 (WLCG und Virtual Observatory legacy)
- SAML (z.B. Helmholtz-AAI, UmbrellaID)
- OIDC (Helmholtz, WLCG und Virtual Observatory migrieren dahin)
- AARC Blueprint Architecture als Basis für viele Implementierungen, was die Kooperation erleichtert

- Authentifizierung außerhalb von Webbrowsern, insbesondere in Cloud-Containern.
- Ggf. Brücken zwischen den verschiedenen Technologien, so dass ein Satz Credentials in einem heterogenen Stack funktioniert.
- Integration der AAI auf Seiten der Serviceprovider.
- Ermöglichung einer fein-granularen (und unter der autonomen Kontrolle der Gruppen oder Kollaborationen bestehenden) Vergabe der Zugangsrechte (→ RDM).

Moderne Instrumente produzieren Daten in Mengen, die auf Arbeitsplatzrechnern nicht zu speichern sind und die verteilten Analysemethoden zugänglich sein müssen.

Die Arbeit mit diesen Daten erfordert föderierte Speichersysteme. Communities sind hier unterschiedlich weit und helfen sich gegenseitig.

Intelligente Storage-Systeme benötigen Echtzeit-Monitoring, um reagieren zu können.

- dCache, XrootD, verteilte Filesysteme (KET, KFS) – skalierbare Storage-Systeme
- ssh, gridftp, webdav/http, Globus (legacy) – Transfer-Protokolle
- Rucio (KET u.a.) – Datenmanagement-System mit globalem Katalog (→ RDM)
- iRODS (z.B. VO) – integriertes Datenmanagement

- Technologien vereinheitlichen und ggf. modernisieren
- Weitere Entwicklungen, um wachsenden Datenvolumina und wachsender Komplexität durch höhere Effizienz entgegenzuwirken
- FAIR-Prinzipien in der Infrastruktur verankern (→ RDM)
- Notwendigkeit für Übertragung *zwischen* Zentren minimieren
- Föderierte Storage-Infrastrukturen (Data Lakes) mit einheitlichem Namensraum und intelligenten Datenplatzierungs- und Replikationsmechanismen
- Dynamisches Buchen von Bandbreite via SDN

- Effiziente Nutzung von nicht-lokalen föderierten und community-übergreifend genutzten Computing-Ressourcen mit transparentem Datenzugriff auf föderierten Speicherinfrastrukturen
- Analyse von großen Datensätzen, die zu groß zum Transferieren und/oder zur lokalen Verarbeitung sind (ggf. im Backend von Science Platforms, → UI)
- Effiziente und nachhaltige Nutzung von opportunistischen Computing-Ressourcen, wie HPC-Zentren oder kommerziellen Cloud-Infrastrukturen

- WLCG Middleware
- COBaID/TARDIS für das Verteilen von Jobs
- AUDITOR für Monitoring und Accounting
- dynamische Datencaches für effizienten Datenzugriff von opportunistischen Ressourcen, z.B. XCache
- Science Platforms/Analysis Facilities, z.B. Jupyter Hubs, Apache Spark, als UIs

- Produktive Installation in großem Stil
- Anpassung an sich stetig ändernde IT-Umgebungen inklusive Neuentwicklungen
- Skalierung
- Grüne Methoden (z.B. Container-Checkpointing), Umgang mit heterogenen, ggf. effizienteren Ressourcen (z.B. GPUs)
- Kopplung mit Analyse-Plattformen und föderierter AAI.
- Effizienter Datenzugriff von opportunistischen Compute-Ressourcen (HPC, kommerzielle Clouds)

In der Schnittmenge zu UIs bewegen sich Science Platforms bzw. Analysis Facilities; aus FI-Sicht sind dabei wichtige Punkte:

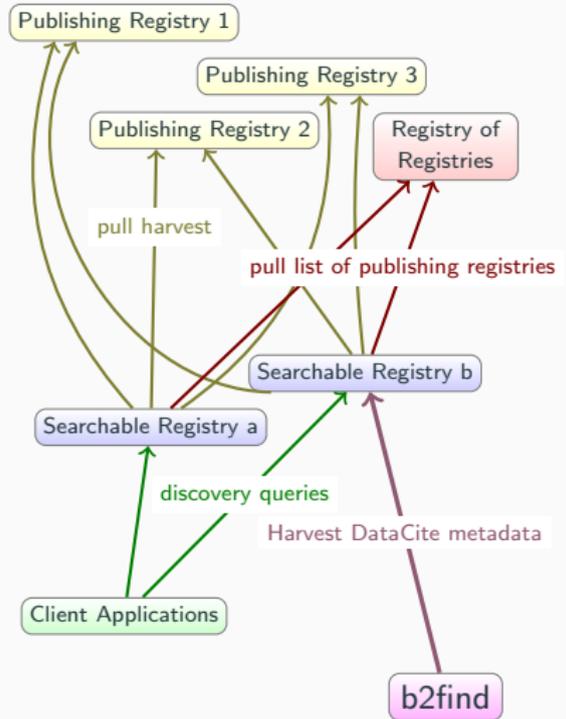
- Schnittstellen zwischen Föderation (v.a. Storage und Computing, aber auch Archiv) und Plattformen
- Unsere Stimme in den Communities, die die großen Lösungen entwickeln
- Plattformen FAIR machen: wie können Skripte und Ergebnisse portabler zwischen Analyseplattformen gemacht werden?
- Skalierung der Lösungen auf Datensätze im Tera- und Petabytebereich

Nach- und Neunutzung bereits genommener Daten spart zumindest Geld und Ressourcen; Beobachtungen in der Astronomie sind i.d.R. ohnehin nicht wiederholbar.

Elemente einer Archiv-Infrastruktur („FAIR“):

- Finden, bevorzugt durch physikalische Beschreibung („blind discovery“)
- Zugriff („Accessible“): Hier besteht häufig eine Schnittstelle zu AAI
- Nutzbarkeit („Interoperable“): Eine neue Datenquelle soll leicht in eigene Workflows integrierbar sein
- Re-usable: Bestehende Arbeiten sollen schnell und einfach nachvollzogen und für eigene Zwecke angepasst werden können.

- Zenodo (CERN) – Zentrale Infrastruktur mit interoperablen nicht-disziplinären Metadaten (DataCite)
- „Virtual Observatory“, globale Data Discovery mit Standard-Diensten und einer zentralen Registry
- Interdisziplinäre Portale wie b2find, openAIRE, EOSC Marketplace



- Global Daten finden durch Definition physikalischer Eigenschaften (Daten-, Experiment-Typ, Empfindlichkeit, beobachtete Positionen, verwendete Stoffe usf.: *Blind Discovery*).
- U.a. dafür: Interoperable Recherche- und Zugriffsprotokolle.
- Disziplinäre Such-Schnittstellen (→ RDM) und ihre interdisziplinäre Förderierung.
- Kuratierung in föderierten Infrastrukturen: Qualitätssicherung (→ RDM), Techniken zur Bereinigung *in der Föderation*.

Immer mehr Geräte produzieren Event-Ströme, die schnell verarbeitet werden müssen (→ RDM, BDA).

Manche davon (z.B. Vera Rubin Telescope) müssen

- global verteilt
- und gefiltert bzw. angereichert werden. Sie müssen am Ende mit hinreichend Information kommen, um ggf.
- rasch über Reaktionen (z.B. Nachbeobachtungen) entscheiden zu können.

- VOEvent, VTP (klassisch)
- Kafka (z.B. bei Rubin/RDS, aber auch als Komponente von dCache/KET)
- AMPEL

- Event Stream Discovery
- Föderation, Gatewaying
- Interoperable Konsumenten (→ UI)

- Dauerhaft finanziert Betrieb laufender Infrastrukturen in sich ständig entwickelnden Umgebungen
- Gesicherte Weiterentwicklung und Pflege der Dienste
- Gemeinsame Förderung von Universitäten und Großforschungseinrichtungen
- Anpassen existierender Lösungsansätze an die Bedarfe weiterer Disziplinen
- Übergang von Prototypen zu Technologien, die die Wissenschaftler:innen nutzen *wollen* (auch: Ausbildung und Support, → KD)
- In bestehende Forschungsumgebungen integrierte Clientsoftware (→ UI)

Zusammenfassung und Ausblick

Vision nachhaltige Rechenzentren

Im Dig-UM Sustainability Paper (arXiv:2311.01169) werden nachhaltige Rechenzentren diskutiert, die grünen Strom nutzen, wo er verfügbar ist.



Techniken zu Föderation und Interoperabilität erlauben, große, ressourcenschonende Infrastrukturen mit der Orientierung auf die Bedürfnisse der Disziplinen zusammenzubringen.

- Schaffung von technologischen Grundlagen, um mit weniger Umweltbelastung mehr und bessere Erkenntnisse zu gewinnen
- Betrieb mit erneuerbaren Energien. Hier gibt es auch spannende Entwicklungen in der Industrie
- Data Lake: Aufbau einer föderierten Storage-Infrastruktur
- Föderation und Nutzbarmachung verteilter heterogener Computing-Infrastrukturen unter Berücksichtigung von Datenlokalität
- Blind Discovery: Auffinden bestehender Daten nach physikalischen Kriterien (z.B. Materialien, Empfindlichkeit)
- So einheitlich wie möglich, so disziplinar wie nötig: Übergreifend verwendbare Komponenten (weiter-) entwickeln und ausrollen



Research Data Management

Astrid Schneidewind
für die DIGUM – Topic – Group

BMBF – Strategiegelgespräch

RDM als ErUM-Thema

RDM ist Thema vieler komplementärer Initiativen – NFDI, HGF, Landesinitiativen, EOSC...
Für effiziente Nutzung der Ressourcen: Abgrenzung und Überlapp definieren.

Hier: **Organisation, Aufbereitung und Bereitstellung der Daten für die Analyse durch Wissenschaftler**

- Statusanalyse, Aktionsplan → Themenbereiche
 1. Entwicklung von (modularen) Systemen zum Forschungsdaten-Management.
 2. Entwicklung von Systemen zur Verwaltung und zum Zugriff auf Metadaten.
 3. Entwicklung von Datenworkflows für effizientere Durchführung von Experimenten und deren Steuerung.
 4. Entwicklung von Systemen für effizientes Life-Cycle-Management.
 5. Entwicklung von Systemen zur Automatisierung und Verwaltung von Zugriffsrechten.

Status in den Communities

KET, KAT, KHuK – Exa-Scale Daten mit hochgradig standardisierter, automatisierter Verwaltung / Bearbeitung

KfB – ebenso mit anderen Raten aber – in-situ

RDS – große Bandbreite an Arbeitsweisen, wie oben bis individuell

KFS, KFN, KFSI – bisher Erzeugung der Daten an Zentren und Verarbeitung in den Forschergruppen; im Wandel durch höhere Datenmengen und –raten, aber noch keine etablierten Strukturen und Verantwortlichkeiten

Status in den Communities

KET, KAT, KHuK – Exa-Scale Daten mit hochgradig standardisierter, automatisierter Verwaltung / Bearbeitung

KfB – ebenso mit anderen Raten aber – in-situ

RDS – große Bandbreite an Arbeitsweisen, wie oben bis individuell

KFS, KFN, KFSI – bisher Erzeugung der Daten an Zentren und Verarbeitung in den Forschergruppen; im Wandel durch höhere Datenmengen und –raten, aber noch keine etablierten Strukturen und Verantwortlichkeiten

High-end Entwicklungen auf der einen Seite, um Fortschritt zu sichern

Entwicklung von angepasstem Datenmanagement an die Bedingungen von Nutzerbetrieb in Zentren unter Übertragung vorhandener Erfahrungen

Internationale Vernetzung der Communities

Verschiedene Arbeitsweisen, Größen der kooperierenden Gruppen, Diversität der Anwendungen

Alle jeweils:

Hochgradige und etablierte internationale Vernetzung mit etablierten Strukturen, inkl. Daten-Management

International genutzte und kontinuierlich gemeinsam entwickelte Analyse- und Steuer- Tools

Nachhaltigkeit - Ressourcennutzung

Verschiedene Arbeitsweisen, Größen der kooperierenden Gruppen, Diversität der Anwendungen

Nachnutzung von Daten / Ergebnissen (inkl. open data).
(Wieder-, Weiter-) Verwendung von vorhandenen Lösungen.
Wissenstransfer – Nachhaltigkeit der Planungen.

Betrachtung der richtigen (Ressourcen-)Einheiten!
Bewusstsein über Impact von Lösungen.

Nachhaltiges RDM ist wichtiger als Neuheit von Lösungen – immer wieder Zugriff auf Daten.
Nur nachhaltige RDM ermöglicht Fortschritt in der Wissenschaft.

Nachhaltigkeit - Ressourcennutzung

Verschiedene Arbeitsweisen, Größen der kooperierenden Gruppen, Diversität der Anwendungen

Nachnutzung von Daten / Ergebnissen (inkl. open data).
(Wieder-, Weiter-) Verwendung von vorhandenen Lösungen.
Wissenstransfer – Nachhaltigkeit der Planungen.

Betrachtung der richtigen (Ressourcen-)Einheiten!
Bewusstsein über Impact von Lösungen.

Nachhaltiges RDM ist wichtiger als Neuheit von Lösungen – immer Zugriff auf gleiche Daten.
Nur nachhaltige RDM ermöglicht Fortschritt in der Wissenschaft.

Resource-aware Research on Universe and Matter:
Call-to-Action in Digital Transformation

Ben Bruers¹, Marilyn Cruces², Markus Demleitner³, Guenter Duckeck⁴,
Michael Düren⁵, Niclas Eich⁶, Torsten Enflin⁷, Johannes Erdmann⁶,
Martin Erdmann^{6*}, Peter Fackeldey⁶, Christian Felder⁸, Benjamin Fischer⁶,
Stefan Fröse⁹, Stefan Funk¹⁰, Martin Gasthuber¹, Andrew Grimshaw¹¹,
Daniela Hadasch^{9,12}, Moritz Hannemann⁸, Alexander Kappes²,
Raphael Kleinemühl¹³, Oleksiy M. Kozlov¹⁴, Thomas Kuhr⁴,
Michael Lupberger¹⁵, Simon Neuhaus¹³, Pardis Niknejadi¹, Judith Reindl¹⁶,
Daniel Schindler¹⁷, Astrid Schneidewind⁸, Frank Schreiber¹⁸,
Markus Schumacher¹⁹, Kilian Schwarz¹, Achim Streit²⁰, R. Florian von Cube²⁰,
Rodney Walker⁴, Cyrus Walther⁹, Sebastian Wozniewski¹⁷, Kai Zhou²¹

1. Entwicklung von (modularen) Systemen

Fokus auf Verwendung und Anpassung von existierenden Systemen oder Integration solcher in wissenschaftliche Lösungen.

- **Standardisierte Lösungen** für Management von **heterogenen Wissenschaftsdaten** (reflektierend die ausgeprägte Diversität innerhalb und zwischen den ErUM-Communities).
- Programme und Interfaces für **leichten Zugriff auf und zum Verwalten** von großen (Meta)Datenmengen, inkl. User-Interfaces, web- oder kommandozeilen-basiert, insbesondere für
 - Umsetzung der FAIR Prinzipien, einheitlicher Zugriff und Auswertung verschiedener Experimente, publizierbare Daten.
 - Katalogisierung / Auffindbarkeit von diversen, in großen Mengen anfallenden Daten.
- **Plattformen für Management** von Trainings- und Validierungsdatensätze **für ML** – communityübergreifend, analog zu Referenz-Datenbanken (→ FI-Storage, BDA).

Beispiel: Rucio – von ATLAS nach cross-community

- Entwickelt für ATLAS, schon von vielen weiteren Communities genutzt.
- Stellt eine ausgereifte, modulare Förderung für RDM bereit.
- Standortbezogen, verwaltet Daten in heterogener, verteilter Umgebung: Erstellung, Lokalisierung, Übertragung, Löschung, Annotation und Zugriff, Orchestrierung von Datenflüssen mit Low-Level und High-Level Richtlinien.
- Frei, open-source Software (lizenziert unter Apache2.0).
- Offener community-getriebener Entwicklungsprozess.
- Option auch für ErUM-Communities.
- Nahtlose Integration von wiss. und kommerziellen Speichersystemen und Netzwerken
- Globaler, einheitlicher Namensraum für die Datenspeicherung
- Verschiedenste Daten möglich
- Nutzer können an geografisch unabhängigen Standorten verteilt sein und zu verschiedenen Verwaltungsbereichen gehören.
- Mehr als 10 Jahre Betriebserfahrung in der Verwaltung sehr großer Datenmengen

Beispiel: Rucio – von ATLAS nach cross-community

- Entwickelt für ATLAS, schon viele weitere Communities.
- Stellt eine ausgereifte, modulare Förderung für RDM bereit.



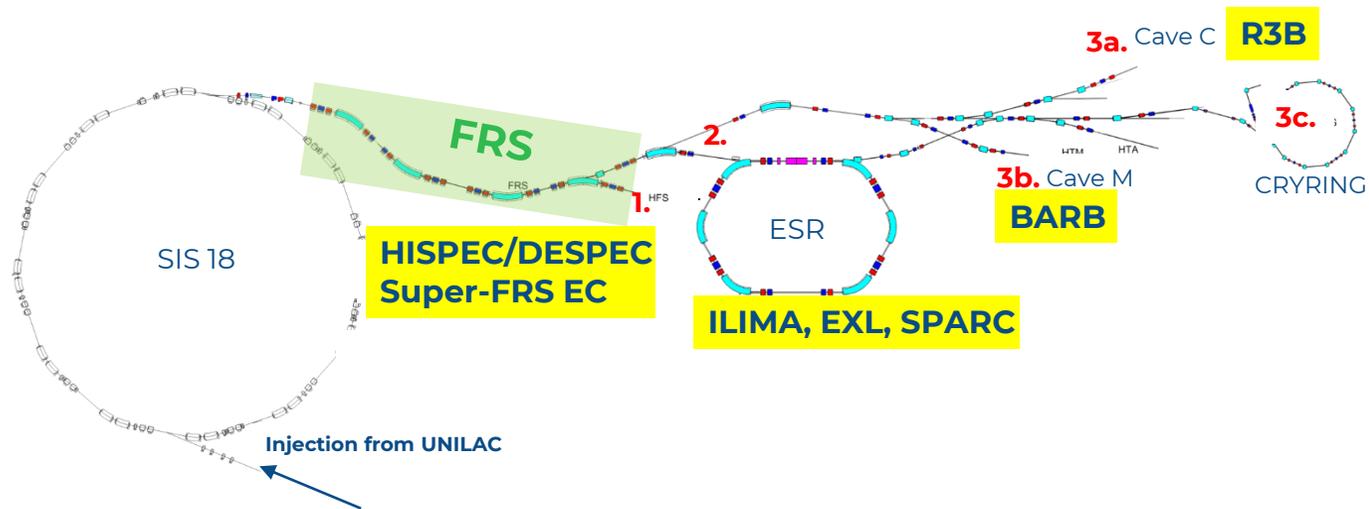
- Nahtlose Integration von wiss. und kommerziellen Speichersystemen und Netzwerken
- Globaler, einheitlicher Namensraum für die Datenspeicherung
- Verschiedenste Nutzdaten möglich
- Nutzer können an geografisch unabhängigen Standorten verteilt sein und zu verschiedenen Verwaltungsbereichen gehören.
- Mehr als 10 Jahre Betriebserfahrung in der Verwaltung sehr großer Datenmengen

2. Systeme für Zugriff auf / Verwaltung von Metadaten

- Modulare, homogenisierte und community-übergreifende Systeme zum **Management und zur Dokumentation der Metadaten sowie Daten-Provenienz**. Metadaten inkludieren Parameter für die Datenreduktion und Datenauswertung. Sicherstellung, dass die einzelnen Schritte von Rohdaten zur Abbildung in der Publikation nachvollziehbar sind. Idealerweise sollten die Metadaten auch (Vor-)Experimente von Laboreinrichtungen beinhalten.
- Workflows zur **flexiblen Gestaltung der Metadatenstruktur** für spezifische Experimente und/oder Instrumente.
- „Blind Discovery“ (→ FI) – Metadatenmanagement dafür nicht skalierbar oder sequentiell.

Beispiel: Multi-used FRS beamline

- Verschiedene Experimente und Nutzerkollaborationen an einer Beamline → einheitliche aber flexible Metadatenstruktur gewünscht.



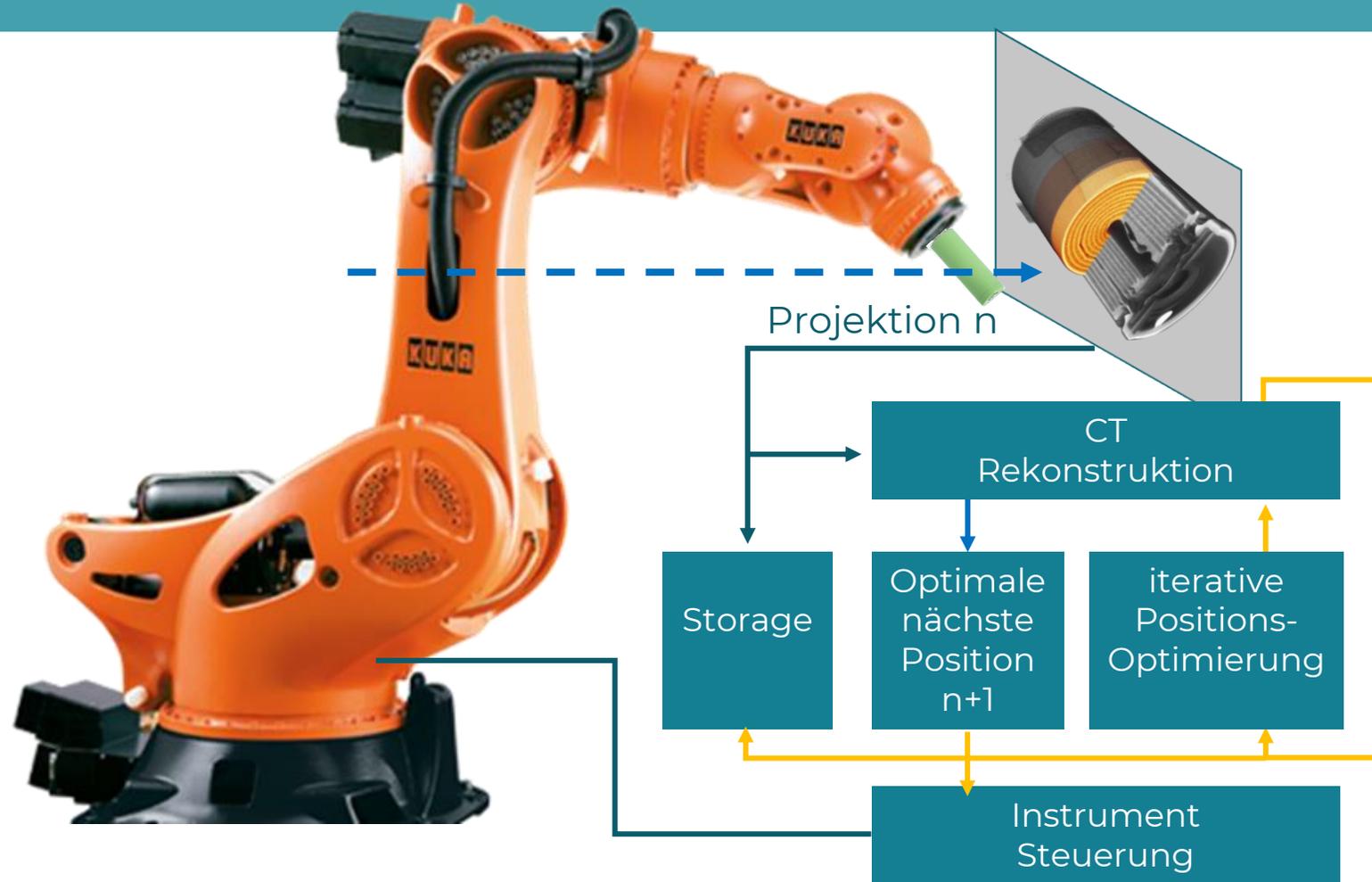
- Versch. spektroskopische, Lebensdauer- und Impulsmessungen
- Reaktionsstudien
- Bio-medizinische Exp. mit e⁺-Emittern
- Astrophysikalische Reaktionen (γ -Fenster)

3. Datenworkflows für effizientere Durchführung von Experimenten und deren Steuerung

- Workflow-Systeme zur **Reduktion von (Roh-)Daten** durch Anwendung von speziellen Algorithmen, z.B. Klassifizierung, Auswahl, und Ausrangieren von Daten sowie die entsprechenden Metadaten, dynamische Metadaten.
- Systeme zur **effizienten Datenkomprimierung** (verlustfrei oder nicht verlustfrei).
- Workflow-Systeme zur Verarbeitung und Auswertung von **Daten ähnlicher Experimente von verschiedenen Großforschungseinrichtungen.**
- Monitoring und Validierung der Daten selbst, Monitoring von Daten-Workflows und der Konsistenz (Datenverlust und Replika Management) von **sehr großen Datenmengen.**
- Workflow Systeme für die effiziente und automatische Visualisierung, Organisation und Extraktion **relevanter Informationen** aus großen Datensätzen, um Experimente zu steuern und fortzuführen.

Beispiel: Roboter-CT

- Effiziente Messzeitnutzung → möglichst wenige, der Probengeometrie angepasste Projektionen zur Rekonstruktion des 3D-Volumens
- Tatsächliche Messpositionen fehlerbehaftet (Genauigkeit) → iterative Optimierung → variable Metadaten
- Bestimmung nächster Messposition auf Basis bisheriger Messungen
- Auswertung und Optimierung in Echtzeit zur Instrumentsteuerung



4. Systeme für effizientes Lifecycle-Management

- Automatisierung von Replika-Management und Strategien zur Optimierung der **Verfügbarkeit relevanter und interessanter** Daten.
- Strategien und Systeme zum **(semi-)automatischen Löschen** von sehr großen Forschungsdaten.
- Strategien und Systeme zur für **optimalen Ausnutzung verfügbarer Computing-Ressourcen** (insbesondere unter Einbeziehung von Nachhaltigkeits-Aspekten) – RDM entwickelt die Modelle und deren Implementierung, FI organisiert und betreibt die Systeme.

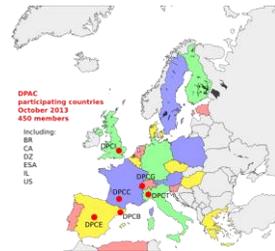
Example: Gaia Data Management

Data Management: Production

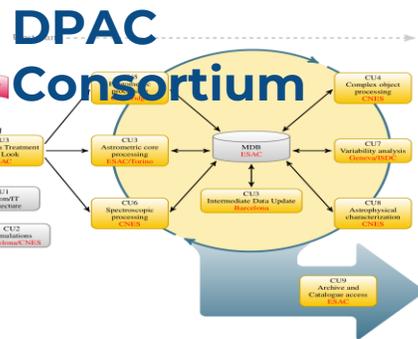
Data Management: User Access



ESA:
Satellite
Operatio
n



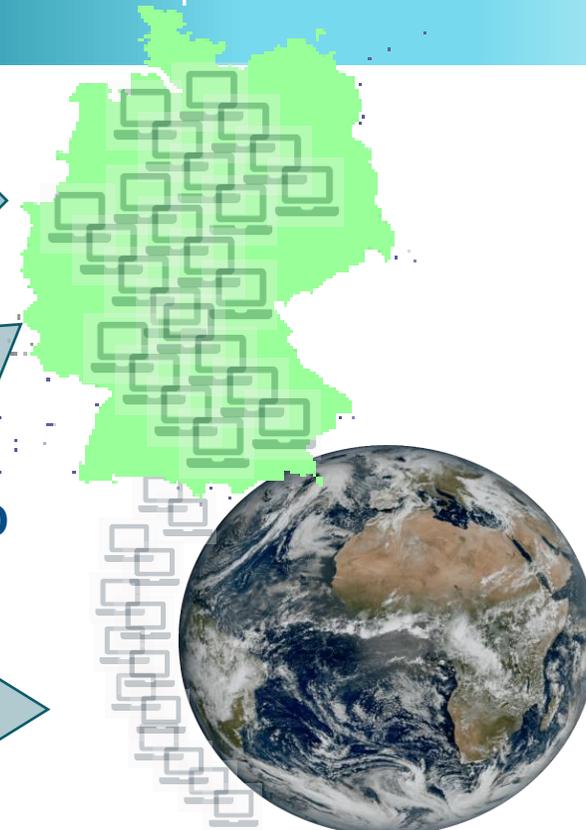
German DPAC participation



Virtual Observatory APIs

10420
journal
papers w.
Gaia data

Data exploitation



Images: ESA + DPAC

5. Systeme zur Automatisierung und Verwaltung von Zugriffsrechten

- Systeme zur Verwaltung von spezifischen (Personen- und Maschinen-) **Zugriffsrechten auf Daten und Metadaten** (von “Private Data” bis “Open Access”). Die Prinzipien werden in NFDI entwickelt, die technische Verwaltung ist aber integrativer Bestandteil der Data- und Workload-Management-Systeme und deshalb technisch in deren Entwicklung einzuschließen.
- Vereinheitlichung und Homogenisierung der Authentifizierung und der Zugriffsrechte auf Forschungsdaten (innerhalb und zwischen den ErUM-Communities).

Take-Home Messages

Nachhaltigkeit ist mehr als CO₂e-Abdruck und grüne Technologien, sondern Effizienz in Bezug auf alle Ressourcen:

- inkl. Personaleinsatz und Wissenstransfer: Arbeiten einmal richtig und nicht mehrfach, parallel oder immer wieder angefangen – deshalb: Nachhaltig = nachhaltige Förderung
- Daten FAIR → mehrfach nutzen, vollständig auswerten
- Workflows FAIR und transparent → Dokumentation erlaubt Nachvollziehbarkeit, damit auch neue Interpretation und neue Erkenntnisse.

Wir stehen damit für „Nachhaltigkeit von Forschung und Umwelt“.

1. Entwicklung von (modularen) Systemen zum Forschungsdaten-Management
2. Entwicklung von Systemen zur Verwaltung und zum Zugriff auf Metadaten
3. Entwicklung von Datenworkflows für effizientere Durchführung von Experimenten und deren Steuerung
4. Entwicklung von Systemen für effizientes Life-Cycle-Management
5. Entwicklung von Systemen zur Automatisierung und Verwaltung von Zugriffsrechten

Vision: Replicate – Reproduce – Re-use



3R

- Große Datenmengen sind verfügbar und nutzbar, weil das RDM es ermöglicht. (Man kann KI nicht ohne RDM denken).
- Standardisierung ist strukturell effizient und erlaubt Reproduzierbarkeit.
- Nutzung existierender Lösungen für viele Bereiche ist nicht nur nachhaltig, sondern auch effizient.
- Wissenschaftler nutzen unsere Systeme gern.
- Wissenschaftler, die sich hier engagieren, werden gewürdigt und haben eine Perspektive in der Wissenschaft. Sie ermöglichen bessere wissenschaftliche Ergebnisse.
- IT-Support wird entsprechend seiner Bedeutung gewürdigt.
- **RDM macht Spitzenforschung nachhaltig erfolgreich.**

Big Data Analytics

Thomas Kuhr
für die DIG-UM Topic Group

BMBF Strategiegelgespräch



Big Data Analytics (BDA) ↔ ErUM Data



Aktionsplan ErUM-Data

Big Data
Analytics
Topic Group



Software
und
Algorithmen

Software und Algorithmen

Software und Algorithmen
entscheidend für ErUM-Wissenschaften
auf international konkurrenzfähigem Niveau



Big Data

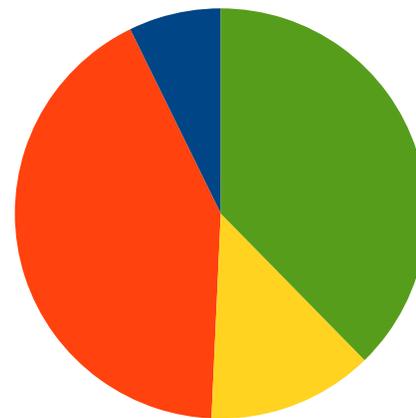
Smart Data

BDA Topic Group

- Umfrage 2022 mit 70 Teilnehmern

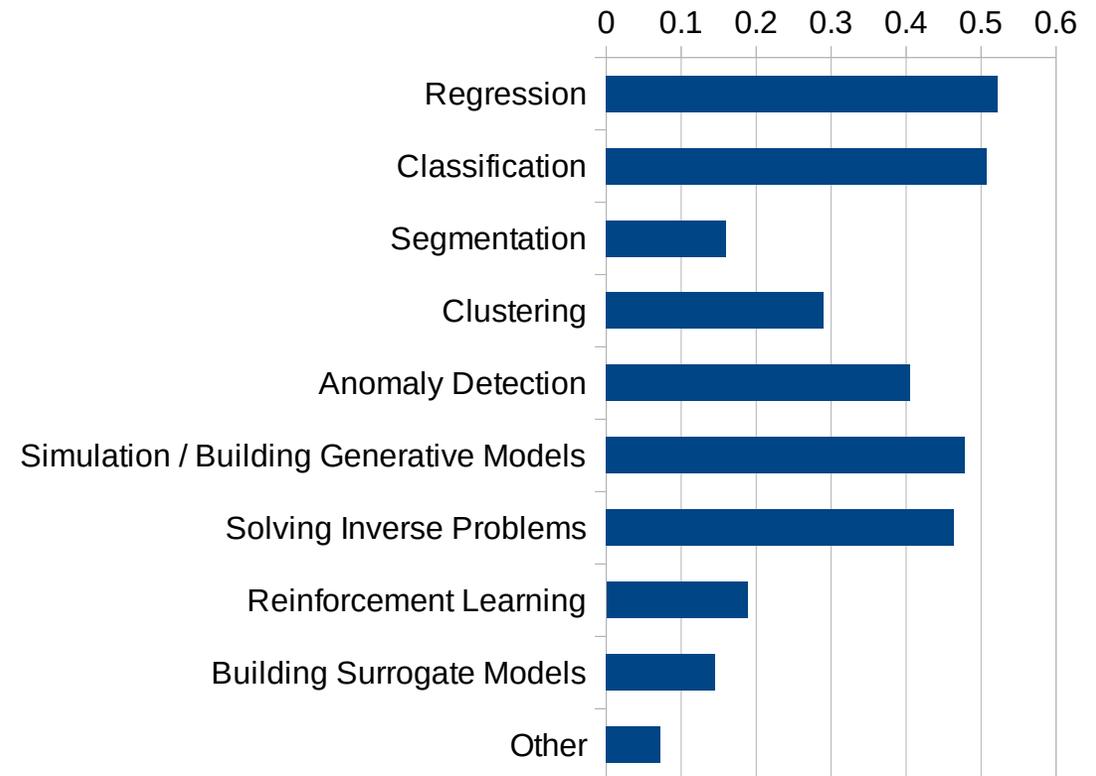
Community	Anzahl Antworten
KAT	10
KET	13
KfB	5
KFN	9
KFS	28
KFSI	0
KHuK	5
RdS	5

What is your overall usage of modern data analysis techniques?



■ No practical usage so far
■ Applied in some projects
■ Heavy user
■ Method Developer

Which tasks are you trying to solve in your research?



BDA Topic Group Organisation

- Offene Mailingliste mit ~60 Subscribern
- Virtuelle Treffen eines Organisations-Teams (Verbundsprecher und Community-Vertreter) etwa alle zwei Monate, Minutes: <https://wiki.erumdatahub.de/de/big-data-analytics>
- Jährliches Treffen
- Workshops zu ausgewählten Themen (Inverse Problems, Generative Models, Realtime Systems)

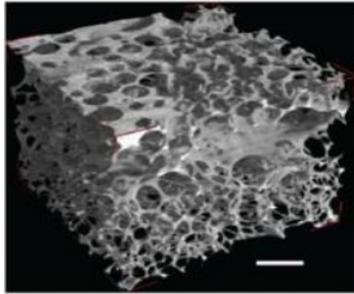
Herausforderungen

- Steigende Datenraten und -volumina (oft mit seltenen Ereignissen)
→ Schnelle und effiziente Algorithmen
- Sparse Data
- Nutzung von modernen Technologien und Methoden, insbesondere KI und Hardware-Technologien
- Nachhaltigkeit von Forschung und Umwelt
→ Optimierung von wissenschaftlichem Output pro Ressourceneinsatz

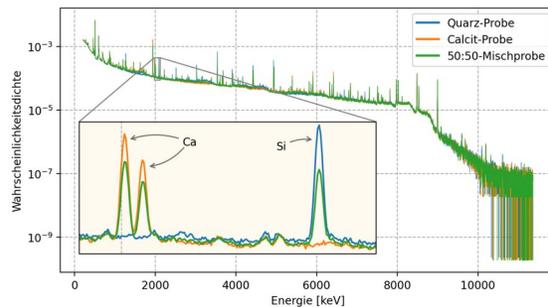
Laufende Projekte

	KAT	KET	KfB	KFN	KFS	KFSI	KHuK	RDS	Inf./M.	Industr.
KI4D4E				X	X				X	X
VIPR		X		X	X					X
KISS	X	X					X	X	X	
ErUM-Wave	X		X						Geoph.	
KI-Morph					X				X	
AI Safety		X							X	X
EvalSpek-ML				X	X		X	X		X
OPAL-FEL			X		X					X
4D-KI-track			X		X					
ErUM-IFT	X					X		X		X

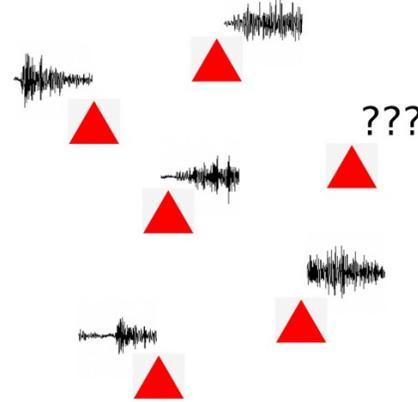
Laufende Projekte



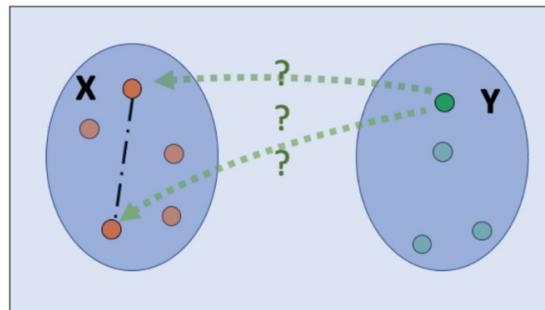
Kompression von 4D Tomographie



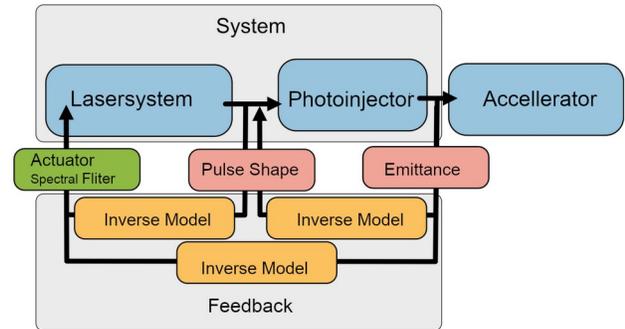
Bestimmung von Komponenten von Spektren



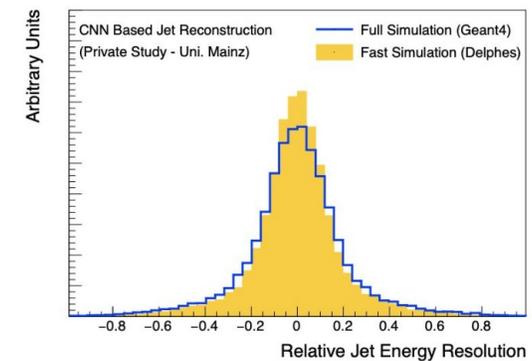
Rekonstruktion von Wellenfeldern



Inverse Probleme



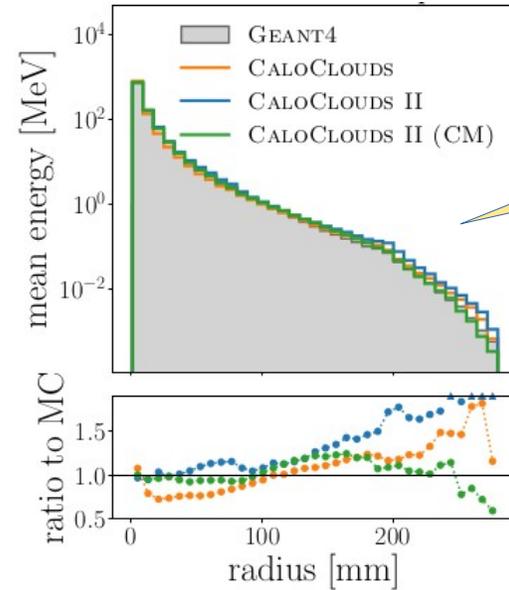
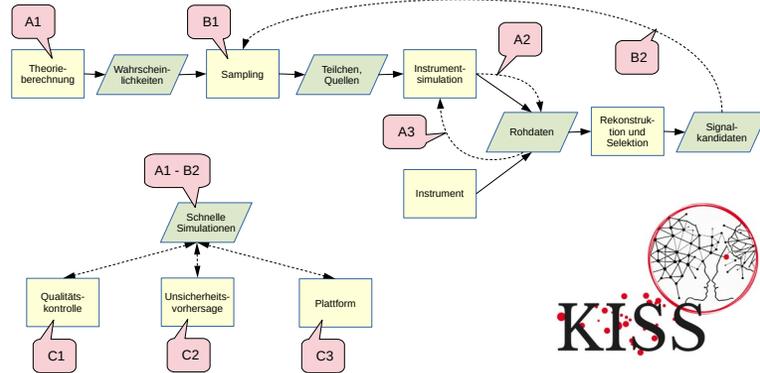
Optimierung von Free Electron Laser



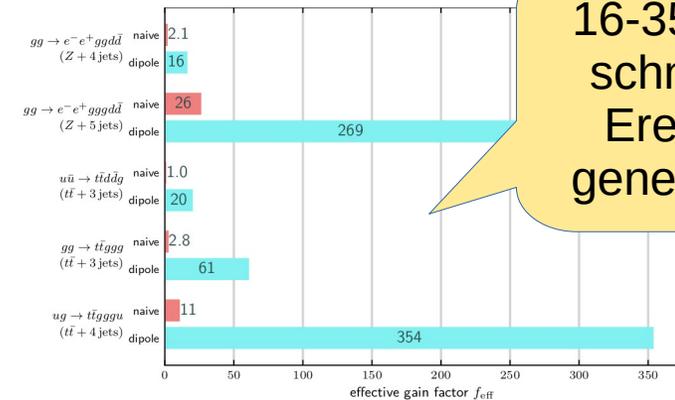
Sicherheit und Robustheit von Neuronalen Netzen

Laufende Projekte: KISS

- **Schnelle Simulationen**



46 mal schnellere Kalorimetersimulation

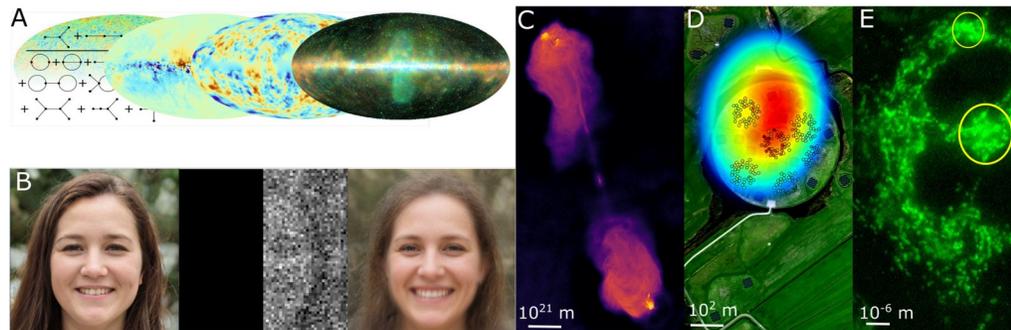


16-350 mal schnellere Ereignis-generierung

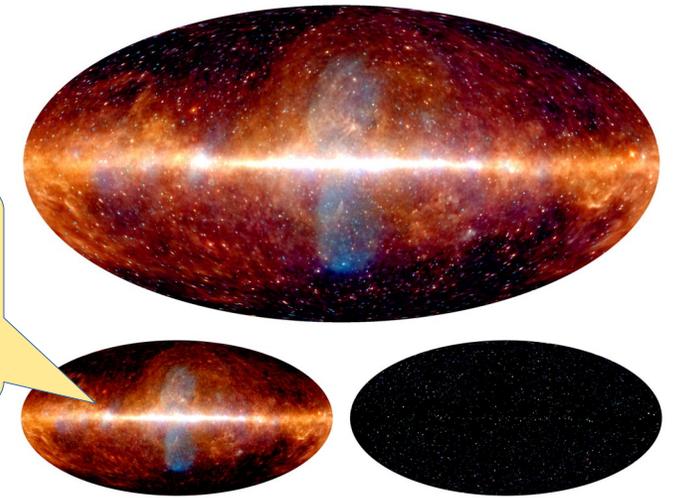
- E. Buhmann, et al, *CaloClouds II: Ultra-Fast Geometry-Independent Highly-Granular Calorimeter Simulation*, arXiv:2309.05704
- T. Janßen, et al, *Unweighting multijet event generation using factorisation-aware neural networks*, SciPost Phys. 15, 107 (2023)

Laufende Projekte: ErUM-IFT

- Information Field Theory Anwendungen



Rekonstruktion von
diffuser Emission
und Punktquellen



- L. I. Platz, et al, *Multi-Component Imaging of the Fermi Gamma-ray Sky in the Spatio-spectral Domain*, *Astron. Astrophys.* 680, A2 (2023)

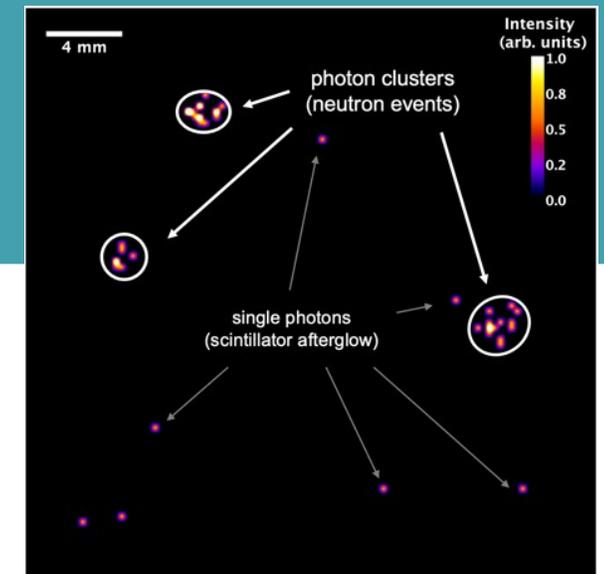
Aktuelle Themen

Beispiele von aktuellen Themen basierend auf

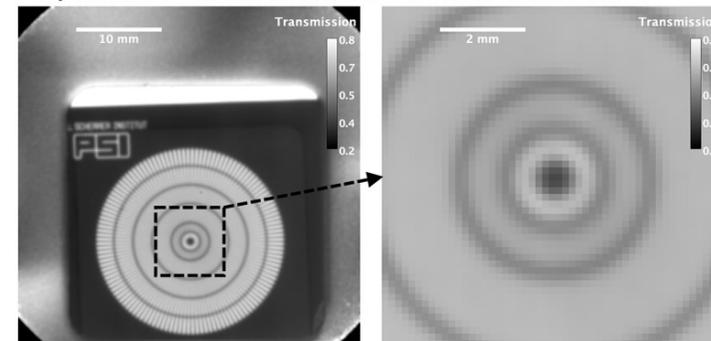
- Umfrage zur Ausschreibung
- Diskussionen auf Jahrestreffen der Komitees
- Individuelles Feedback
- Diskussion in BDA Organisationsteam

Realtime Algorithmen

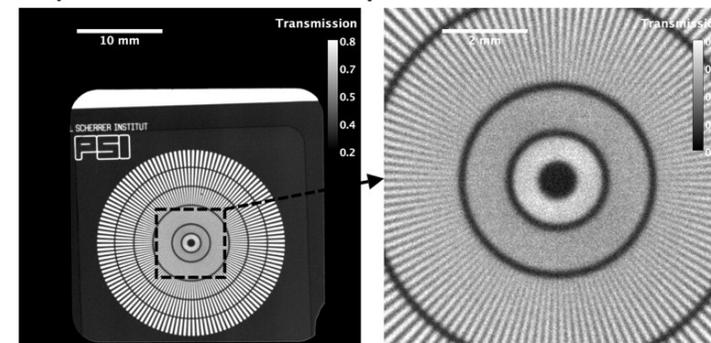
- Herausforderung: Steigende Datenraten
- Schnelle Entscheidungen erforderlich für Online-Analyse/Selektion, Erkennung von Anomalien und Experiment-Optimierung (→ RDM)
- ➔ Entwicklung/Optimierung von schnellen Algorithmen für (heterogene) Systeme mit CPUs, GPUs, FPGAs, etc.



A) Photon event mode at native detector resolution



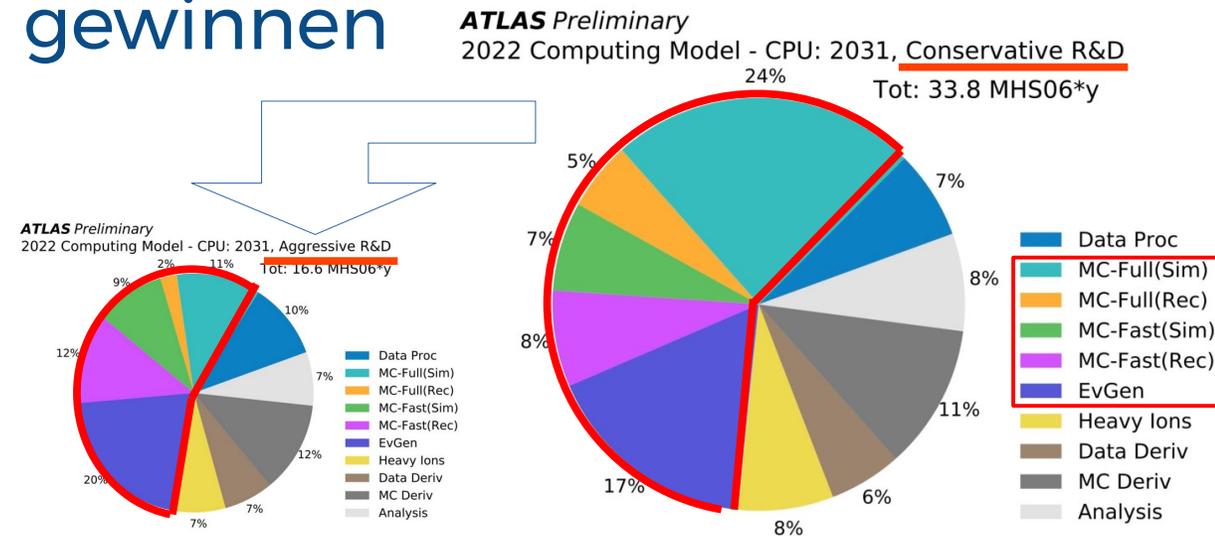
B) Neutron event mode at super resolution



Sci Rep. 2021 Nov 1;
11(1):21360

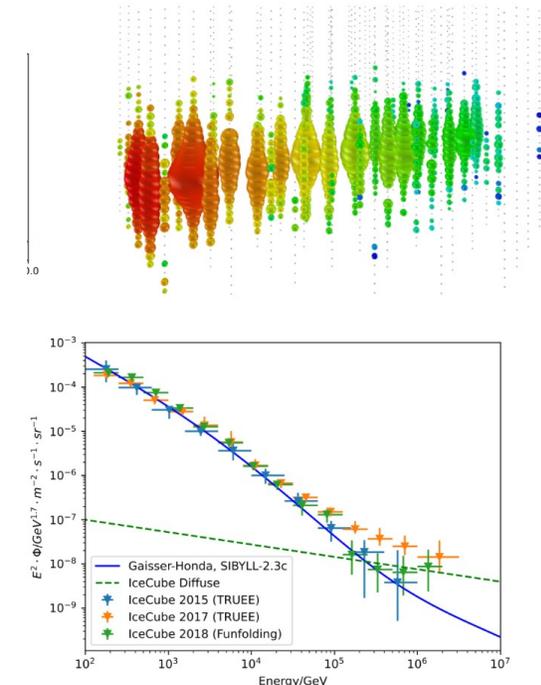
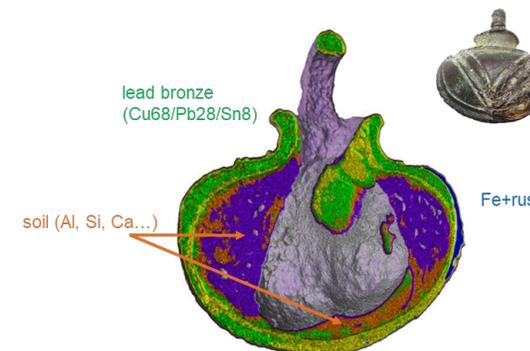
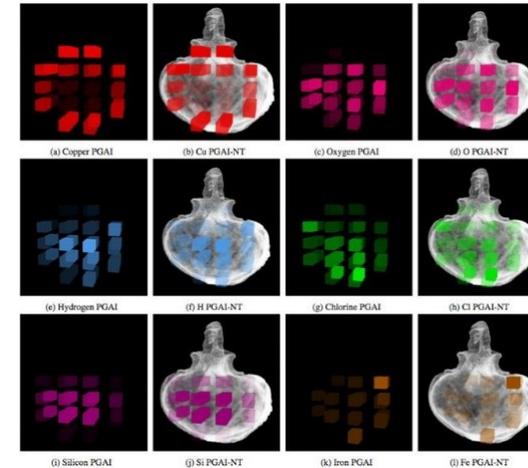
Generative Modelle

- Herausforderung: Steigender Ressourcenbedarf
- Großer Anteil an Ressourcen benötigt für Simulationen, die oft der einzige Weg sind, um wissenschaftliche Ergebnisse aus den aufgezeichneten Daten zu gewinnen
- ➔ Entwicklung und Anwendung innovativer Methoden (KI) zur Beschleunigung von Simulationen



Inverse Probleme

- Herausforderung: Bessere Ergebnisse durch neue Methoden
- Gewinnung von Erkenntnissen über den erzeugenden Prozess aus Daten ein Problem in vielen Wissenschaftsbereichen, insbesondere bei sparse data
- ➔ Erschließung neuer Methoden und Nutzung von Synergien in gemeinsamen Projekten



Nachhaltige Software

- Herausforderung: Effizienz und Nachhaltigkeit von Entwicklungen
- ➔ Experimentübergreifende Entwicklungen
 - Einige Tools werden bereits erfolgreich von mehreren Experimenten genutzt, z.B. GEANT4 für Detektorsimulation
 - ➔ Erweiterung auf weitere Bereiche, z.B. ACTS für Spurrekonstruktion
- ➔ Langfristiger Support, Weiterentwicklung existierender Software
 - ➔ qualifiziertes Personal

Comput Softw Big Sci 6, 8 (2022)

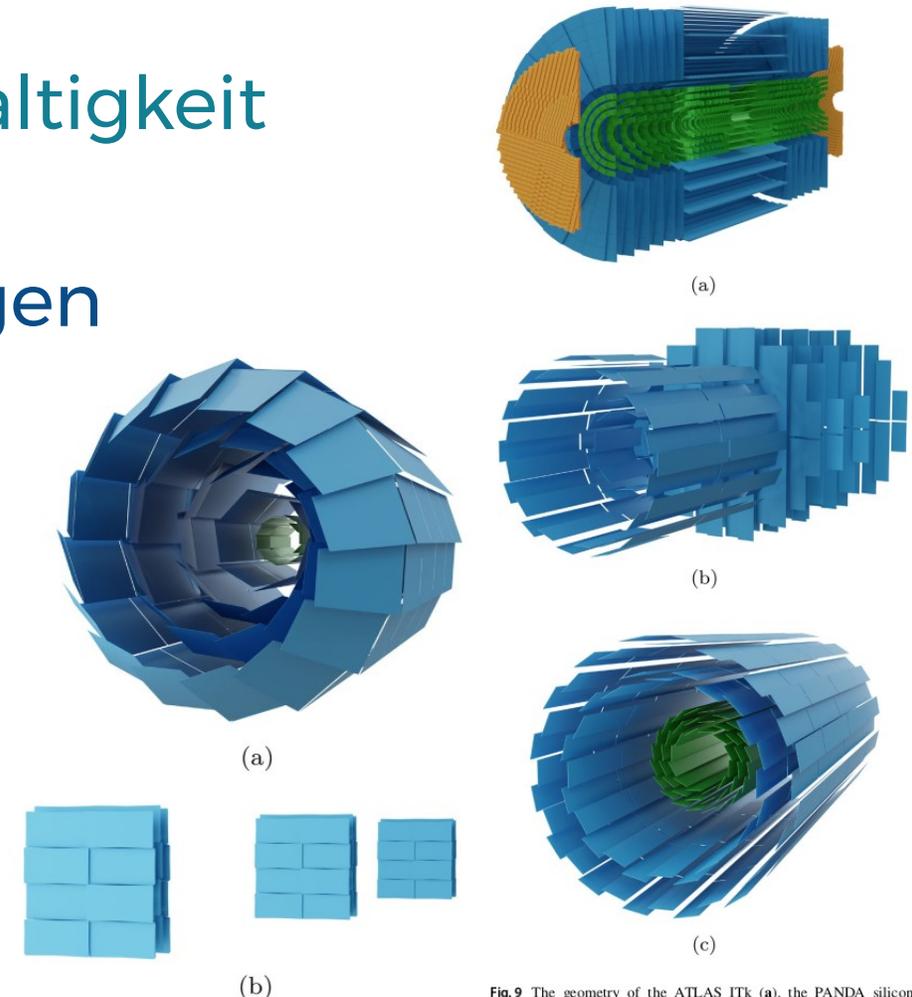


Fig. 10 Geometries of Belle II (a) and FASER (b) implemented in ACTS. Colors indicate the different subsystems

Fig. 9 The geometry of the ATLAS ITk (a), the PANDA silicon detector (b), and the SPHENIX silicon tracking detectors (c), implemented with ACTS. Colors indicate the different subsystems; in the top image, the High Granularity Timing Detector (HGTD) [83] is shown in orange

Software-Optimierungen

- Herausforderung: Steigender Ressourcenbedarf, Nachhaltigkeit
- Bessere Auflösung von Instrumenten führt zu größeren Datenmengen
- ➔ Nutzung von modernen Methoden der Software-Entwicklung zur Reduzierung des Ressourcenbedarfs der Datenverarbeitung
- ➔ Anpassung an dynamische Verfügbarkeit von Ressourcen (→ FI)

Resource-aware Research on Universe and Matter:
Call-to-Action in Digital Transformation

Ben Bruers¹, Marilyn Cruces², Markus Demleitner³, Guenter Duckeck⁴, Michael Düren⁵, Niclas Eich⁶, Torsten Enflin⁷, Johannes Erdmann⁶, Martin Erdmann^{6*}, Peter Fackeldey⁶, Christian Felder⁸, Benjamin Fischer⁶, Stefan Fröse⁹, Stefan Funk¹⁰, Martin Gasthuber¹, Andrew Grimshaw¹¹, Daniela Hadasch^{9,12}, Moritz Hannemann⁸, Alexander Kappes², Raphael Kleinemühl¹³, Oleksiy M. Kozlov¹⁴, Thomas Kuhr⁴, Michael Lupberger¹⁵, Simon Neuhaus¹³, Pardis Niknejadi¹, Judith Reindl¹⁶, Daniel Schindler¹⁷, Astrid Schneidewind⁸, Frank Schreiber¹⁸, Markus Schumacher¹⁹, Kilian Schwarz¹, Achim Streit²⁰, R. Florian von Cube²⁰, Rodney Walker⁴, Cyrus Walther⁹, Sebastian Wozniowski¹⁷, Kai Zhou²¹

¹Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany.

²Max Planck Institute for Radio Astronomy, Bonn, Germany.

³Universität Heidelberg, Heidelberg, Germany.

⁴Ludwig-Maximilians-Universität München, München, Germany.

⁵Justus Liebig University Gießen, Gießen, Germany.

⁶RWTH Aachen University, Aachen, Germany.

⁷Max Planck Institute for Astrophysics, Garching, Germany.

⁸Forschungszentrum Jülich GmbH, Garching, Germany.

⁹TU Dortmund University, Dortmund, Germany.

¹⁰Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany.

¹¹University of Virginia, Charlottesville, Virginia, USA.

¹²University of Tokyo, Kashiwa, Japan.

¹³Bergische Universität Wuppertal, Wuppertal, Germany.

¹⁴Heidelberg Institute for Theoretical Studies, Heidelberg, Germany.

¹⁵University of Bonn, Bonn, Germany.

¹⁶Universität der Bundeswehr München, Neubiberg, Germany.

¹⁷Georg-August-Universität Göttingen, Göttingen, Germany.

¹⁸University of Tübingen, Tübingen, Germany.

¹⁹Albert-Ludwigs-Universität Freiburg, Freiburg, Germany.

²⁰Karlsruhe Institute of Technology, Karlsruhe, Germany.

²¹Frankfurt Institute for Advanced Studies, Frankfurt, Germany.

*Corresponding author(s). E-mail(s): erdmam@physik.rwth-aachen.de;

arXiv:2311.01169v1 [physics.comp-ph] 2 Nov 2023

Weitere Themen

- Foundation Models
- Modellierung von Zeitreihen
- Anwendungen der Informationsfeldtheorie
- KI-unterstützte Software-Entwicklung
- Quantencomputer-Algorithmen
- ...

Diskussionspunkte

- Explorative Forschung und langfristige Nutzbarmachung von Entwicklungen (→ Nachhaltigkeit) beides gewünscht
→ Auch Weiterentwicklung von existierender Software
- Neue Projekte und Projekte, die auf Entwicklungen aus der ersten Förderperiode aufbauen
- Rechtzeitige Bewilligung hilfreich für Findung/Haltung von Personal
- Transfer wird begrüßt und durch ErUM-Data unterstützt, projektabhängiges Transferpotential zu:
 - a) Anderen Experimenten in derselben Community
 - b) Anderen Communities
 - c) Wirtschaft

Diskussionspunkte

- Erforderliche Beiträge zu Experimenten mit Transferpotential in ErUM-Data statt ErUM-Pro
- Bezug zu ErUM-Transfer?
- Internationale Einbindung wichtig
- Nachhaltigkeit des Fußabdrucks (und der entwickelten Lösungen) wird als sehr wichtig betrachtet
- Flexibilität in der Gestaltung der Verbände
- Begutachtung mit Expertise in Methoden und ErUM-Wissenschaften gewünscht

Zusammenfassung

- **Steigender Bedarf** an Entwicklungen im Bereich Software und Algorithmen zur Bewältigung der Herausforderungen großer Datenmengen, neuer Methoden und Nachhaltigkeit
- **Breites Spektrum an Themen** (breiter als in erster Förderperiode)
- **Transferpotential** (zu anderen Experimenten, anderen Communities, Wirtschaft) kann in der zweiten Förderperiode ausgebaut werden, ist aber projektabhängig
- **Flexibilität** bei Thema und Zusammensetzung von Verbänden für eine bedarfsgerechte Förderung

Vision: Schneller, Nachhaltiger, Mächtiger

- Entwicklung **nachhaltiger Software**
 - für eine **effiziente Nutzung von Daten**
 - unter Verwendung **moderner und etablierter Technologien**
 - zur Optimierung des **wissenschaftlichen Erkenntnisgewinns**
 - im Hinblick auf den **Ressourceneinsatz**
- ➔ **Der für international konkurrenzfähige ErUM-Forschung notwendige Fortschritt in Breite und Tiefe der Algorithmen kann nur mit Unterstützung des BMBF erzielt werden**

The User Interface Topic Group

Ein interdisziplinäres Nutzerportal

Pierre Schnizer, Tim Ruhe et al.
für DIGUM Topic Group User Interface



User Interface: innerhalb von DIG-UM

Was soll gespeichert werden?

Was?

Research Data
Management

Big Data Analytics

Auf welche Weise?

Welche Algorithmen werden
verwandt?

Wie werden die
Daten verwandt?

Wie?

User
Interface

Wo soll es gespeichert werden?

Wo?

Federated
Infrastructures

Knowledge Distribution

Wer?

Wer benötigt das Wissen ?

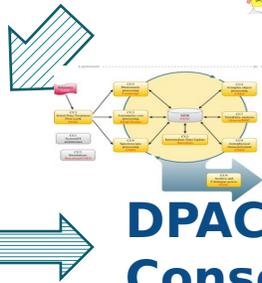
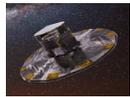
User Interface: Datenfluss

GAIA: Daten auf arbeitung: Zur Verfügung stellen

Analyse



ESA:
Satellite
Operation



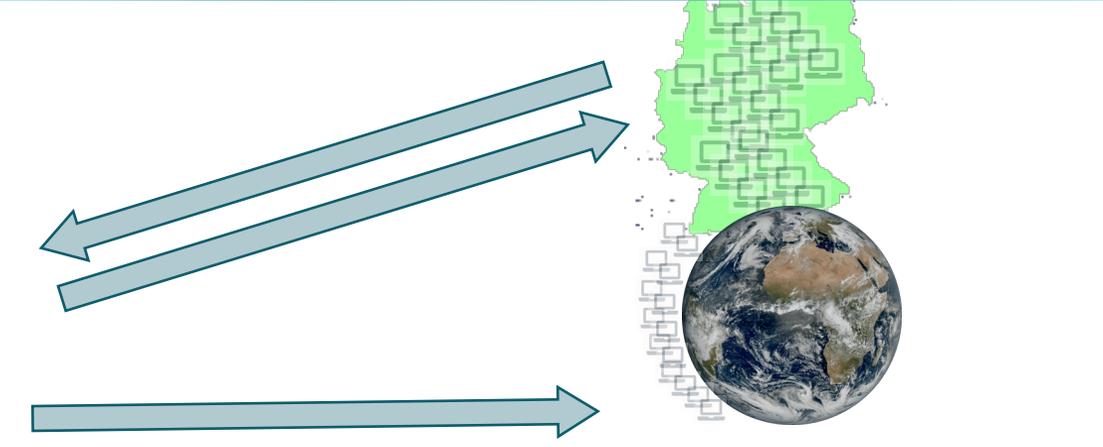
**DPAC
Consortium**

**DPAC+
(Partner)
DataCenters**
:



Images: ESA + DPAC

Gute Lösungen:
Werden hier übernommen



UI ↔ Federated infrastructure: Zugang zum Server: Client

UI ↔ Research data management:

- * Layout Daten / Metadaten

- * Workflow

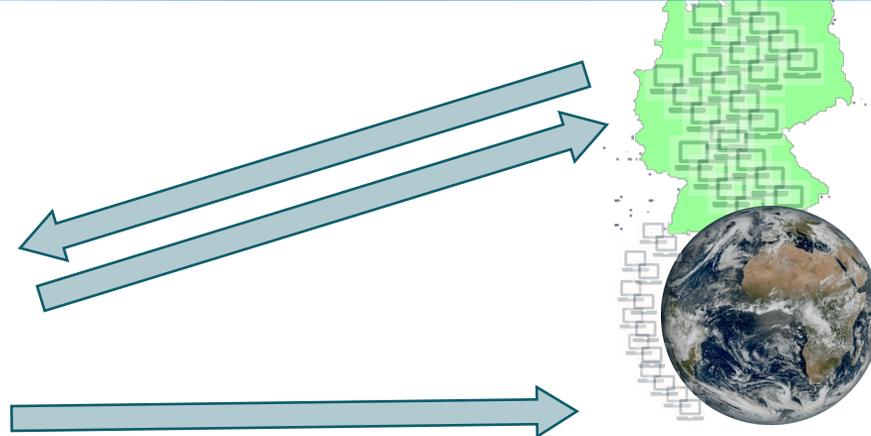
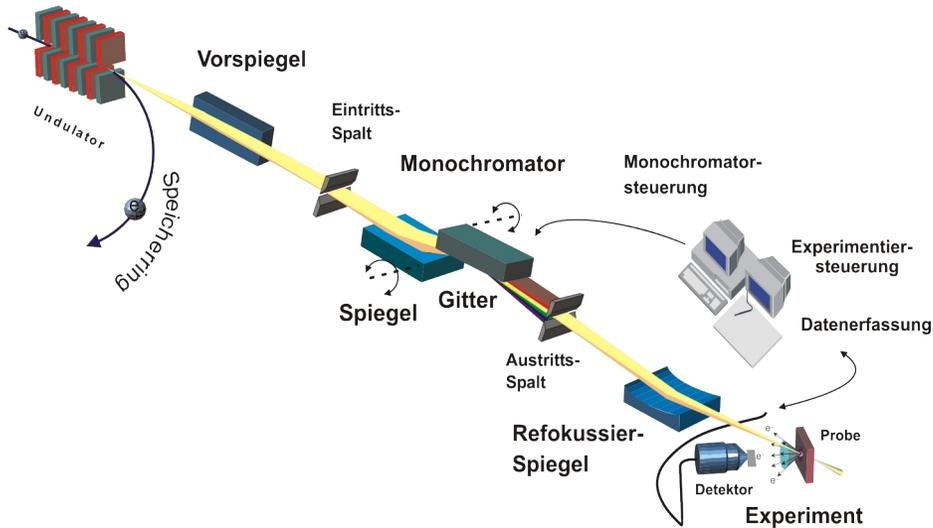
UI ↔ Big data analytics: Algorithmenschnittstelle

Outreach: spread the info

User Interface: Datenfluss

GAIA: Daten auf arbeitung: Zur Verfügung stellen

Analyse



Gute Lösungen:
Werden hier übernommen



UI ↔ Federated infrastructure: Zugang zum Server: Client

UI ↔ Research data management:

- * Layout Daten / Metadaten

- * Workflow

UI ↔ Big data analytics: Algorithmenschnittstelle

Outreach: spread the info

P. Schnizer T. Ruhe et al.

2023: Workshop der Forschungsgemeinschaften



Experten Workshop „Next Generation Environment for Interoperable Data Analysis“

3-4 Mai 2023, HZB Berlin

Sprecher:

- Nicolas Eich, VISPA, RWTH
- Verana Kain, CERN
- Mohammad Al-Turany, GSI / FAIR
- Kai Polsterer, HITS, Heidelberg

User Interface: Ziel

Ideales User Interface (Anforderungen)

- Daten
 - Suchen: so einfach wie mit Google
 - Laden: so leicht wie ein „pip install“
- Analyse
 - Wissenschaftliche Bearbeitung wie ein Nachmittagsspaziergang
 - Support von Workflows, Daten einfach und passend bereitstellen
 - Arbeitsumgebung für
 - wiederkehrende Aufgaben in der wissenschaftliche Bearbeitung

→ Fokus der Wissensschaffenden auf die Wissenschaft

Innerhalb DIG-UM

- Wissenmultiplikation und Transfer zwischen den Forscher:innengemeinschaften
- Unterstützung in Designmuster und Richtlinienentwicklung
- → “Testsammlung“ für wissenschaftliche User Interfaces



[Dieter Rams: gutes Design](#)

Workshop Teilnehmer: Intuitives Dateninterface

„Higher Order Data Products“

- Ähnlich zu Datenkatalogen
- Daten und Analyseketten
- Versionierung:
 - Ursprungsdaten
 - Analysepipeline
- Motivation:
 - flache Lernkurve
 - Basierend auf bewährten Mustern

Workflow / Ablaufplan: Unterstützung

- Validiert
- Dokumentiert
- Archiviert

Workshop Teilnehmer: allgemeine Empfehlungen

Anerkennung: Datenkuratierung

- Motivation:
 - Wissenschaftlicher Mehrwert → Sichtbarkeit
- Wertschätzung für
 - Sorgfältige Datennahme, Speicherung
 - Erstellung und Pflege der Analysepipelines
 - Datendokumentation
- Lösungsvorschlag
 - DOI → Review → zitierfähig

Offene Daten

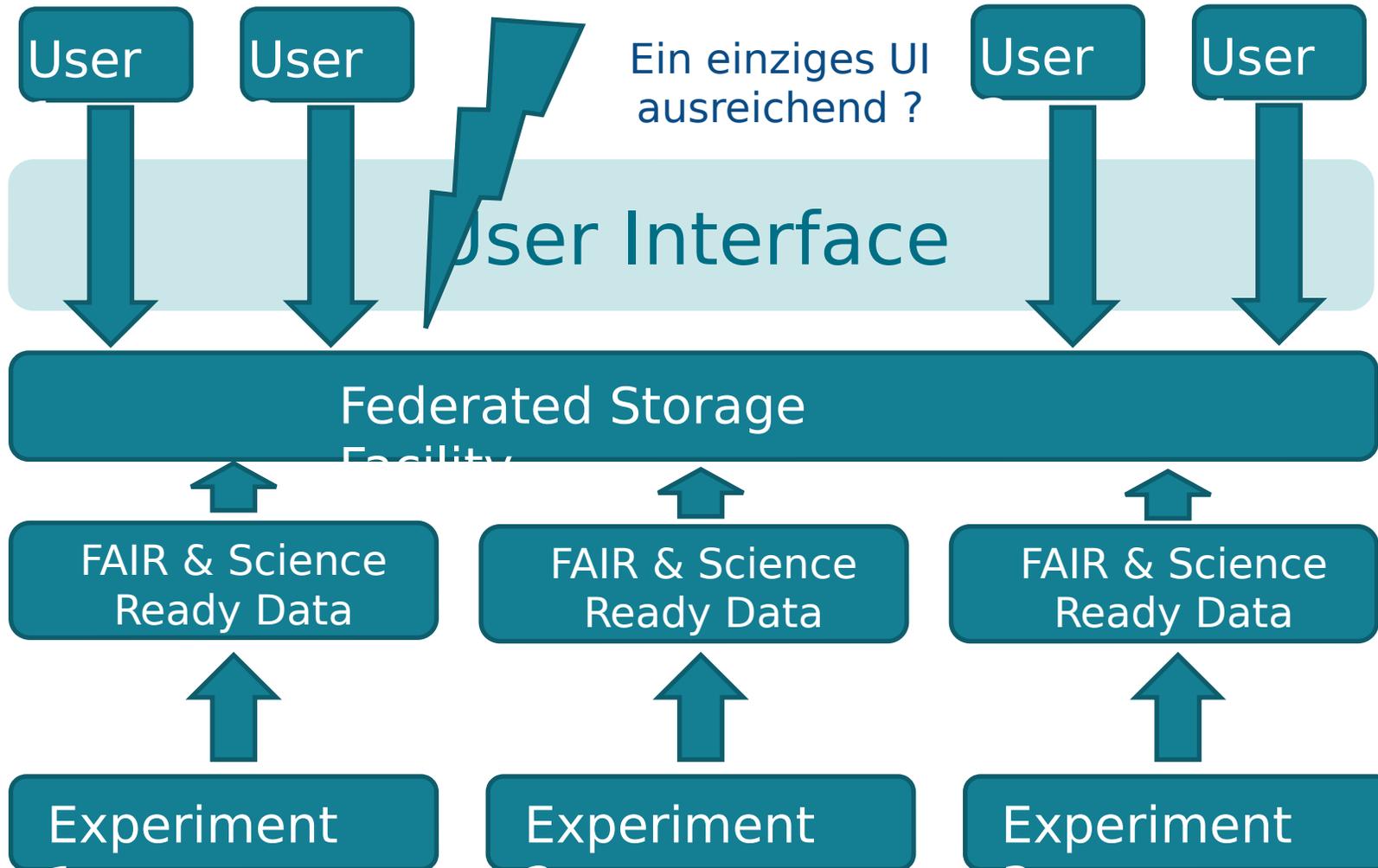
Durch niederschweligen Zugang

- → häufig benutzt
- → verstanden
- → gepflegt

Wissensbasar

- Kommunikationskanal
- Offen zugänglich / bekannt
- → Skaleneffekt

Interdisziplinäre Entdeckungen



Intuitive
Benutzeroberflächen sind
selbsterklärend

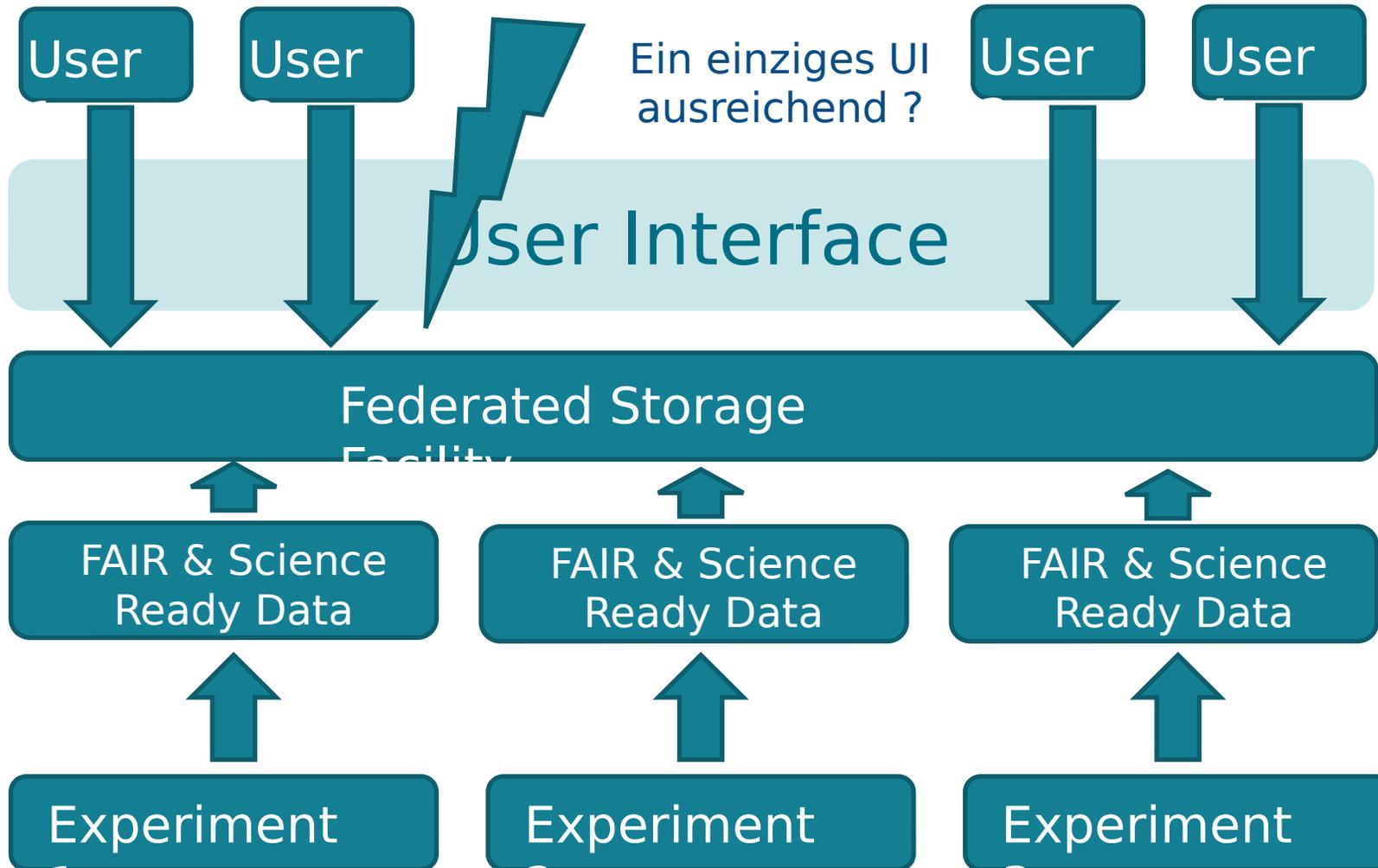
*Deren Design bedeutet
sorgfältige Überlegungen*

Speicherung und Kuratierung
der Daten durch die
Experimente

Zeitnahe Publikation der
Daten durch die Experimente

Verbindung von Datenquellen
verschiedener Forschungs-
gruppen durch gemeinsame
Standards

Interdisziplinäre Entdeckungen



Datenbearbeitung und Kuratierung durch Experimente

Für große internationale Kollaboration möglich, herausfordernd für kleine Forscher:innengruppen

Bereitschaft und Ausbildung notwendig!

Datenzugang: „Ende des Anfangs“

Workflows / Ablaufpläne

Darstellung, Archivierung, Validierung!

Wiederholbare und überprüfbare Workflows
(Programmier-) Sprachen unabhängig

Sorgfältig dokumentiert

Workshop Beispiele: Vispa, Alpha

Erfolgreiches Praxisbeispiel:

Reana (Transfer HEP → Astro)

Graphische Programmierung ?

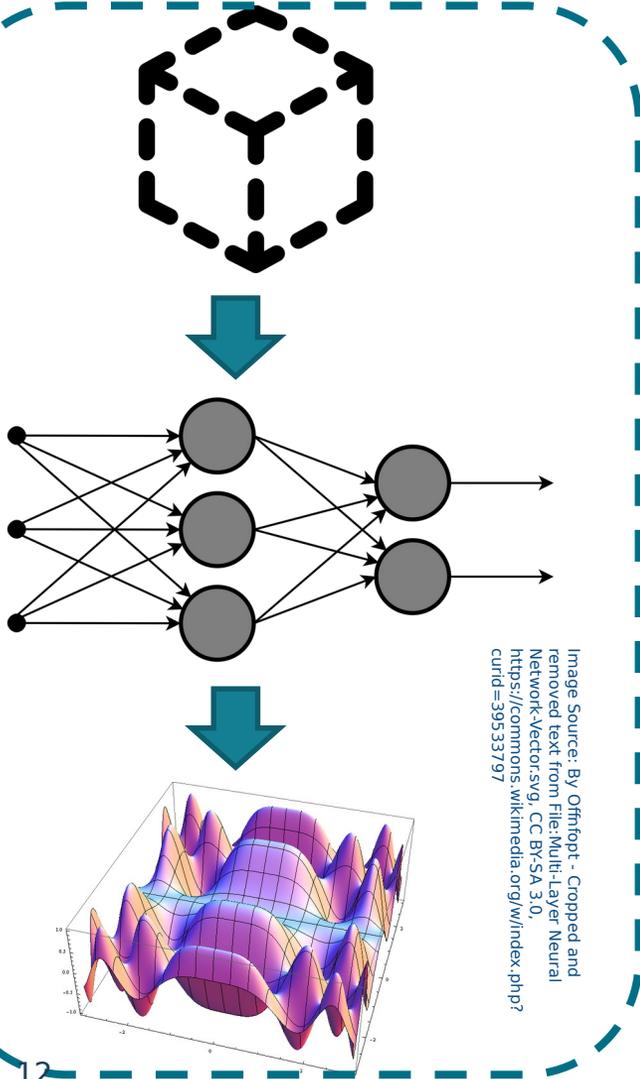
„Low Code“ Ansätze (Konfiguration statt Programmierung)

KI-unterstützte Programmierumgebungen

Teillösungen für viele Probleme existieren

Vereinfachung:

Zugang zu existierenden Lösungen, Vereinfachung ihrer
Kombination



Aufbauend auf Workflows und Daten

Wissensbasar



Image Source: By MCruz (WMF) - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=34370213>

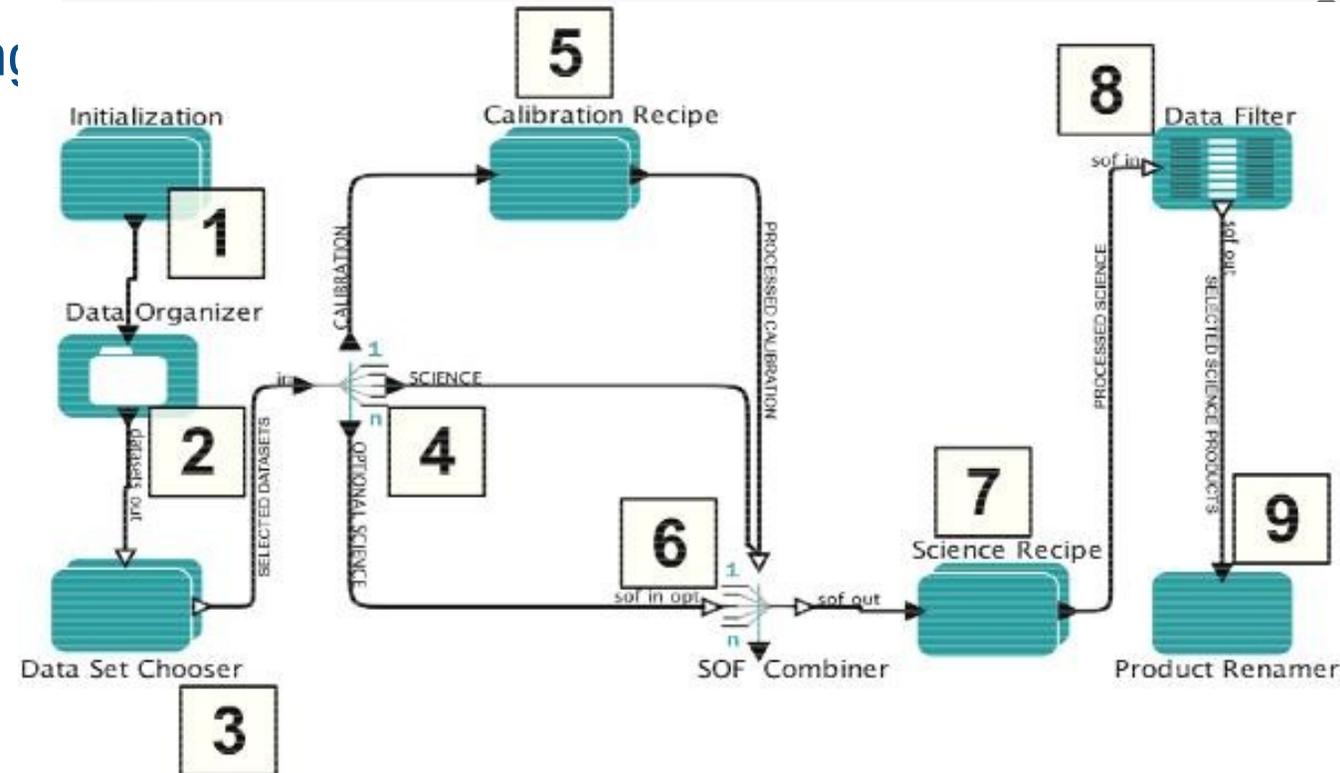
Wissensmanagement
unabhängig von
einzelnen Personen

Peer Programming/Co-Pilot
Verwendung von
existierendem Code und
existierenden Workflows
Vorschläge für
wiederverwendbaren
Code / „best practices“
Softwaredesignmuster und
Architektur
Semiautomatische
Erstellung von Code



ChatGPT
Physics

- Datenverarbeitung: Analyse / Validierung
- Komplexe Abläufe
 - Beschreibungssprache
 - User Interface → Vereinfachung: “Konfiguration statt Programmierung“ für wiederkehrende Fälle



[Kepler analysis: flow](#)

- Datenverarbeitung: Analyse / Validierung
- Komplexe Abläufe
 - Beschreibungssprache
 - User Interface → Vereinfachung: “Konfiguration statt Programmierung“ für wiederkehrende Fälle

Lösungsansätze

- Unterstützung der Entwicklungsumgebung (ähnlich wie e.g. Pycharm / Python type annotation)
- Low code Platform
- Graphische Programmierung

- große Projekte / Anbieter:
 - Weboberflächen angeboten
 - wissenschaftliche Analysen mit komplexer Software

e.g:

- Jupyter Hub → Parallelisierung: Dask
- Apache Spark: Analyse Engine
- Workflow Beschreibung

- Schnittstellen Föderation ↔ Plattformen
- Platform discovery:
 - Daten am Ort X
 - Werkzeug Y
 - Autorisierung z
 - → möglicher Anbieter
- “FAIR”e Plattformen: Austausch → Skripte, Ergebnisse
- → Skalierung und Performanz

Erarbeitung / Sammlung von Richtlinien

- Gemeinsame Anforderungen
 - Designempfehlungen: z.B.: Synchrotron Lichtquellen Beamline Interface
 - Datenkataloge: Richtlinien für „science ready data“
 - Layout
 - bewährte Muster
 - Standards der Forschungsgemeinschaften
 - Beschreibende Referenzgrafiken
- Daher → „higher order data products“

Vereinheitlichung ← aus den Forschungsgemeinschaften heraus

Unterstützende künstliche Intelligenz

- Ähnlich Co-Pilot / ChatGPT
- Aber: basierend auf
 - logischen Strukturen → Vorhersagen
 - publizierten wissenschaftlichen Texten
- Basierend auf öffentlich zugänglichen Modellen

Abstraktion / Schichtenmodell

- Aufteilung komplexer Interfaces: Designmuster / Architektur
- Für Subsysteme und deren Nahtstellen
z.B: Datenportale, Kataloge, Workflows, nachhaltige Datenprodukte

Langfristige Ziele: Dataprovisionsinfrastruktur

Nachhaltige Daten: Voraussetzung langfristige Förderung

Infrastruktur: Anforderungsprofil

- → kuratierte Daten
 - Notwendig
 - Übernahme von Daten und Metadaten
 - Qualitätsüberprüfung
- Konformitätsprüfung: Standards der Forschungsgemeinschaften
- Speicherung & Zugänglichmachen
 - ggf. Berücksichtigung von Sperrfristen

Innerhalb von DIG-UM

- → Anforderungsprofil
- Richtlinien: „science ready data“
- Sammlung Standard von Metadaten
- Anforderungen an
 - Kuratoren
 - Archivare

Zeithorizont

2022/2023

- Forschungsgemeinschaften erster Erfahrungsaustausch
- Identifikation der Unterstützer und Interessenvertreter
- Erste Bestimmung der Bedürfnisse der Forschungsgemeinschaften
- Erster Workshop

2024/2025/2026

- Ausarbeitung der gemeinsamen Bedürfnisse
- Ableitung der Entwicklungsziele
- Definitionen von Standards für Daten und Metadaten
- Identifikation von „bewährter Praxis“
- Implementierung erster Prototypen

2027/2028/2029

- User Interface(s) Implementierung
- Schrittweise Implementierung weiterer Eigenschaften und Ausbaustufen

Ziele: Fokussierung

- Wissensaustausch: mittels jährlicher Workshops
- Daten- und Metadaten-Standardisierung, Unterstützung in der Entwicklung
- Erforschung und Erarbeitung erreichbarer Ziele
- technische Herausforderungen
- benötigte Zeit
- Förderung bewährter Workflow-Modelle sowie deren Verbreitung in andere Wissenschaftsgebiete
- Nachhaltigkeit der Lösungen
- Wissensaustausch durch praktische Erfahrung

Zweiter Expertenworkshop

Second Workshop on Next Generation Environment for Interoperable Data Analysis

Dortmund Q2/2024

DIG-UM User Interface: Take home message

User Interface:

- gutes Design essentiell:
Intuitiv, selbsterklärend
- Entwicklung des User Interface für
 - hochwertige Datenprodukte
 - Workflows
 - Algorithmen
- für große Datenmengen

Sorgfältige Entwicklung: Basis guten Designs

Motivation: Effizienzsteigerung
Nachhaltigkeit der Entwicklungen
Besser Datennutzung

- Bisher:
 - Aufbau der Vernetzung
 - erster Austausch zwischen Forscher:innengemeinschaften
- Nächste Schritte
 - Ableitung: Bedürfnisse, Entwicklungsziele, Richtlinien
 - Implementierung: Prototypen
- Langfristig
 - UI Implementierung: Schrittweise Vertiefung & Ausbaustufen

DIG-UM User Interface: Vision

User Interface: mehr als Bedienungsoberflächen

- Große wissenschaftliche Datensätze verwenden
- Einfaches bequemes Design der den Datensätze zugrunde liegenden Modelle
- Einfach *wieder* zu verwenden
- Algorithmen: konsistentes, intuitives Design

- Wissensmultiplikation: Richtlinien und Kommunikation

- Data Provision Infrastruktur:
- Übernahme, Kuratierung der Daten,
- bereitgestellt durch intuitives User Interface



<https://www.flaticon.com/free-icons/knowledge-sharing>



Knowledge Distribution

Judith Reindl (KfSI)

Katrin Link (KAT)

Dirk Lützenkirchen-Hecht (KFS)

+ ErUM-Data-Hub:

Angela Warkentin

Peter Fackeldey

Benjamin Fischer

Judith Steinfeld

Ulla Lardinoix

Martin Erdmann



Topic Group Knowledge Distribution & ErUM-Data-Hub



Kontakte & Netzwerk innerhalb und für ErUM Communities

Langfristige Ziele, Inhalte und Orientierung (Nachhaltigkeit)

Strategische Beratung

Berichterstattung

Konzeption, Planung und Durchführung von Maßnahmen

(Workshops / Schulen / Messen / ...)

Haushaltsplanung & Verwaltung / Administration

Evaluation & Dokumentation

Zahlen & Daten



Retrospektive / Start in 2022

Thema mit großer Wichtigkeit und hohem Interesse für ErUM Communities

⇒ **Deep learning / machine learning**

Strategie:

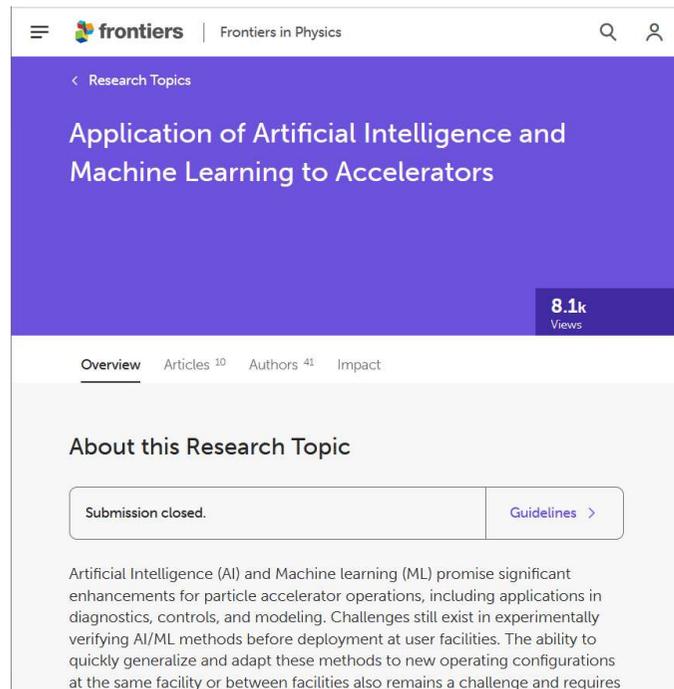
Phase 1: 2-tägige Train-the-Trainer workshops

Ziel:

Befähige TeilnehmerInnen eigenständig Kurse zu Deep Learning zu geben

Phase 2:

- Initiiere eine Serie von Schulen für interessierte Wissenschaftler, Studenten, Doktorierende, ...
- Initiiere eine Serie von Experten-Workshops zu speziellen Themen (RDM, Big Data, Data Sustainability, ...



The screenshot shows a web page from Frontiers in Physics. The main heading is "Application of Artificial Intelligence and Machine Learning to Accelerators" with a view count of 8.1k. Below the heading, there are tabs for "Overview", "Articles 10", "Authors 41", and "Impact". A section titled "About this Research Topic" contains a "Submission closed" button and a "Guidelines" link. The text below discusses the promise of AI and ML for particle accelerator operations and the challenges of experimental verification.



Retrospektive: 3/2022 - 12/2023

- 4 Train-the-Trainer Workshops: Deep-Learning Basics / Advanced
(Aachen, Wuppertal, Dortmund & München)
- 5 Deep-Learning-Schools (Basics / Advanced) – Meinerzhagen
- 6 Experten – Workshops zu verschiedenen Themen: u.a. BDA, RDM, UI, Inverse Probleme, ...
(Wiehl, Bonn, Hamburg, Berlin, Garching, ...)

⇒ In Summe 15 Veranstaltungen mit über 520 Teilnehmenden / 100 % Weiterempfehlung



Techn. Support etc.
durch ErUM-Data-Hub



⇒ Veranstaltungen sind mehr als nur reine Wissensvermittlung:
Netzwerken, Bedarfsermittlung, Wahrnehmung in (wiss.) Öffentlichkeit, ...

Retrospektive: 3/2022 - 12/2023



Σ 15 Events mit >520 Teilnehmenden

5 Deep-Learning-Schulen

6 Experten-Workshops

4 Train-the-Trainer Workshops

100% Weiterempfehlung



Detaillierte Evaluation der Veranstaltungen

Beispielhaft anhand von zwei Experten-Workshops:

Big Data Analytics - Connection Workshop

Feb 23 – 24, 2023
DESY
Europe/Berlin timezone

Overview

Timetable

Contribution List

Registration

This short in-person workshop in Hamburg serves to connect **PIs and project partners of research projects in the ErUM-Data call 'Software & Algorithms'**.

The goal is two find synergies between funded projects (and other activities such as the NFDI consortia DAPHNE and PUNCH) and to discuss how projects that received no funding in this round could still be realised.

The workshop will take place on 23.2. noon to 24.2. on the DESY campus in Hamburg.

ErUM-Data Workshop on Inverse Problems

Dec 5 – 6, 2023
Campus Garching
Europe/Berlin timezone

Overview

Timetable

Registration

Surveys

Payment

Venue & Travel

Accommodation

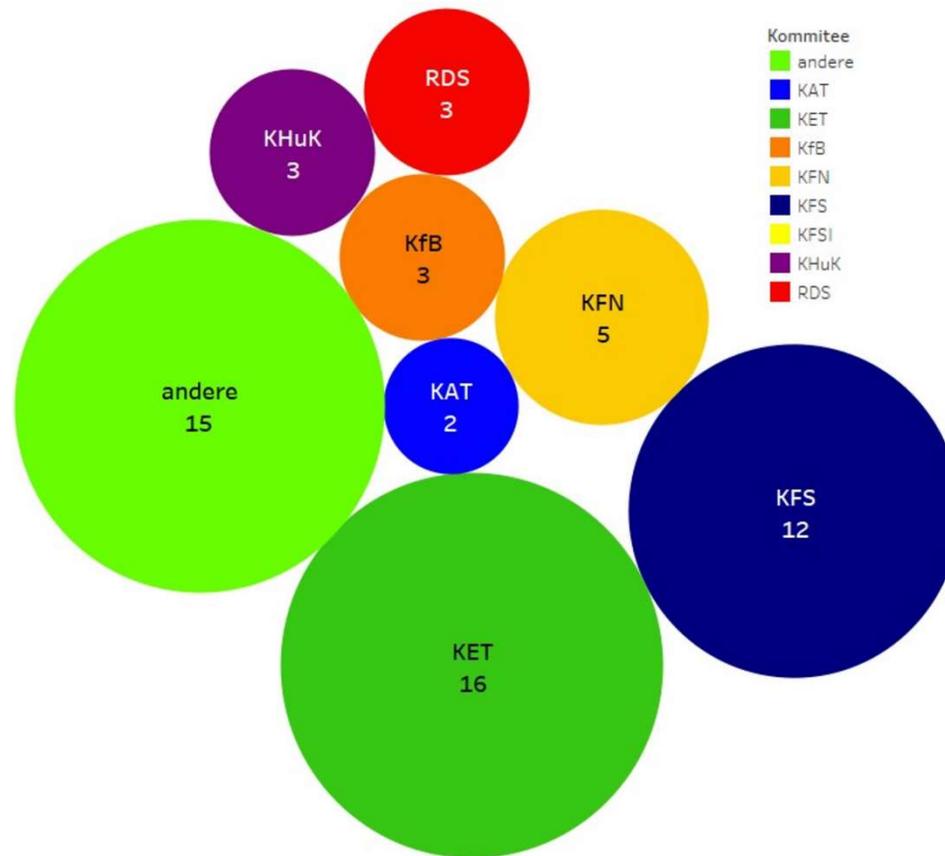
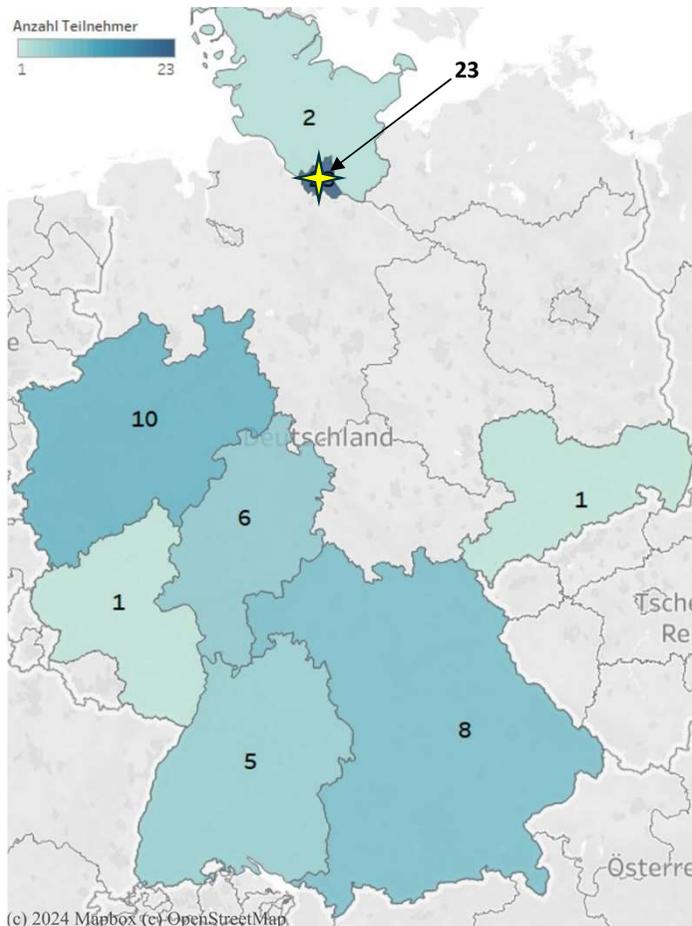
Due to the difficult travel conditions, the workshop is also being broadcasted online. Join here: <https://rwth.zoom-x.de/j/67466745362>

This Workshop is hosted by the community organization **DIG-UM** with support from the BMBF-funded **ErUM-Data-Hub**.

The goal of the workshop is to get an overview on inverse problems in the **ErUM-Data community**, to discuss methods to solve them, and to foster exchange and cooperation between the ErUM-communities. In particular, the workshop is intended for early stage researchers.

Topics to be discussed include:

Expert Workshop Hamburg: Big Data Analytics 2/2023

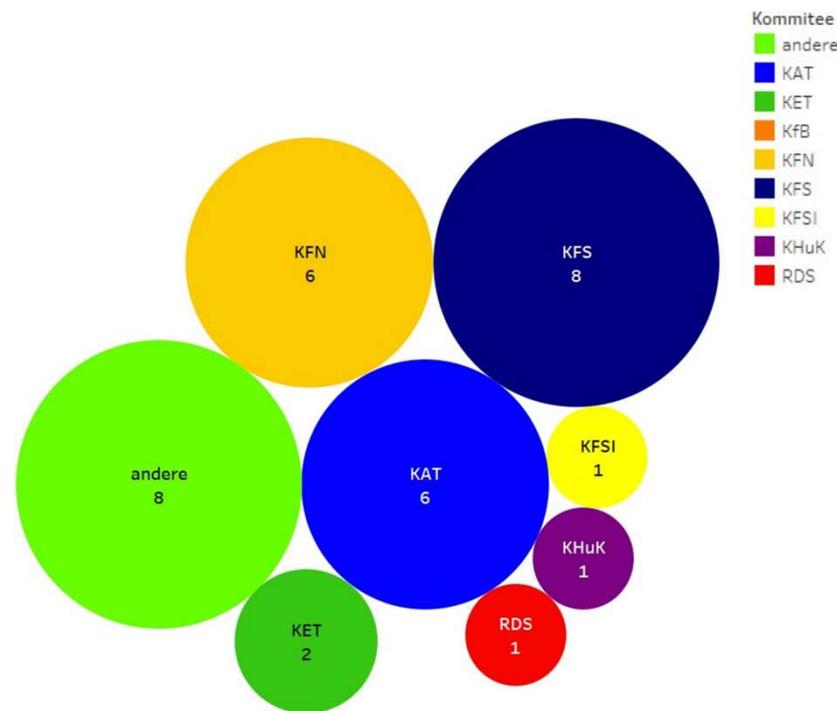
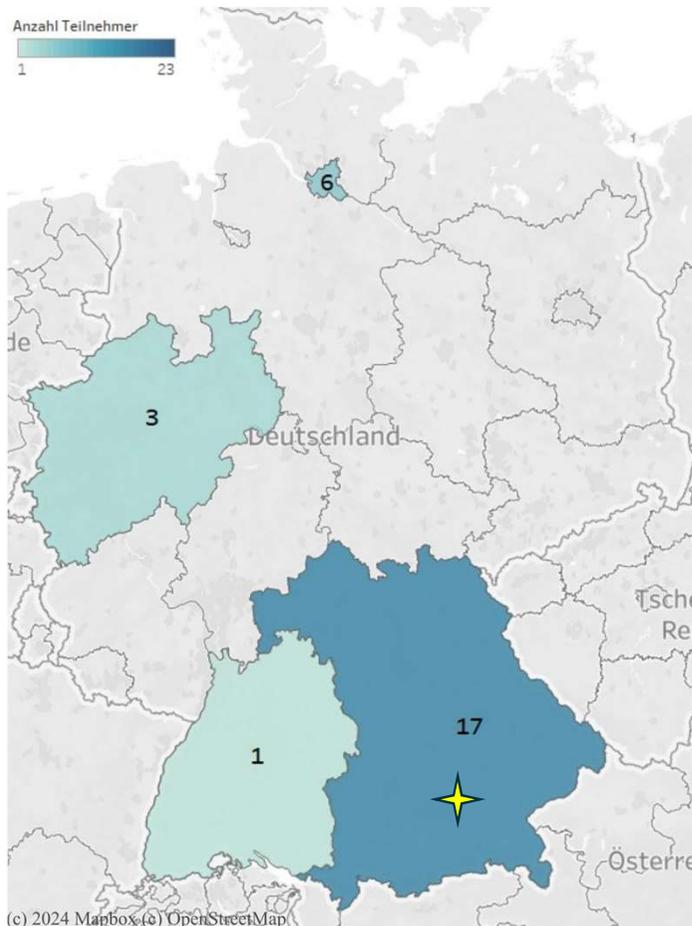


Σ 58 Teilnehmende

Fast 50% aus Hamburg

KET / KFS gut vertreten

Expert Workshop München: Inverse Problems 12/2023

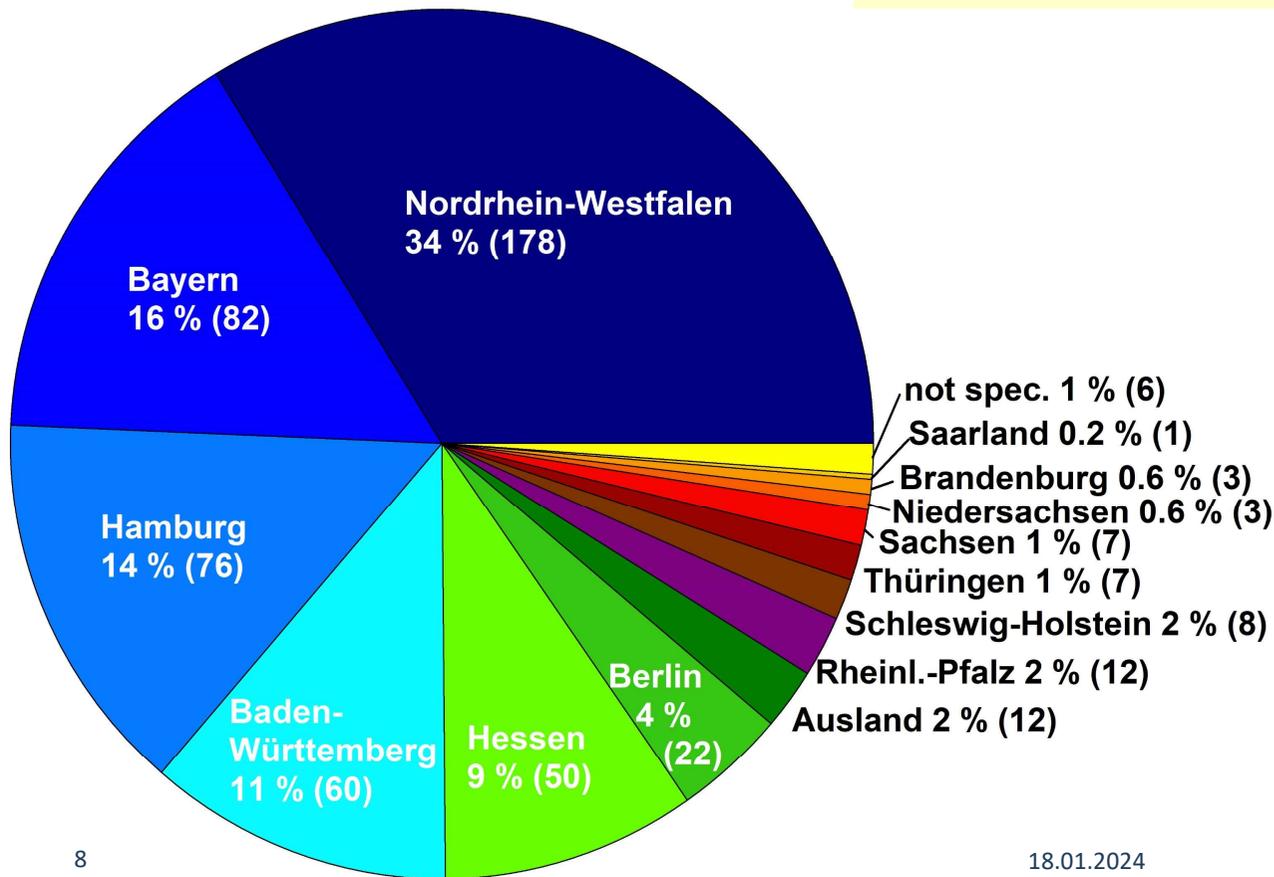


Σ **28 Teilnehmende**
 \approx 60% aus Bayern
KFS, KFN & KAT gut vertreten

Retrospektive: 3/2022 - 12/2023

Evaluation aller Veranstaltungen:

Teilnehmerinnen & Teilnehmer meist aus näherer Umgebung



- Nordrhein-Westfalen (ca. 34 %)
- Bayern (ca. 16 %)
- Hamburg (ca. 14 %)
- Baden-Württemberg (ca. 11 %)
- Hessen (ca. 9 %)

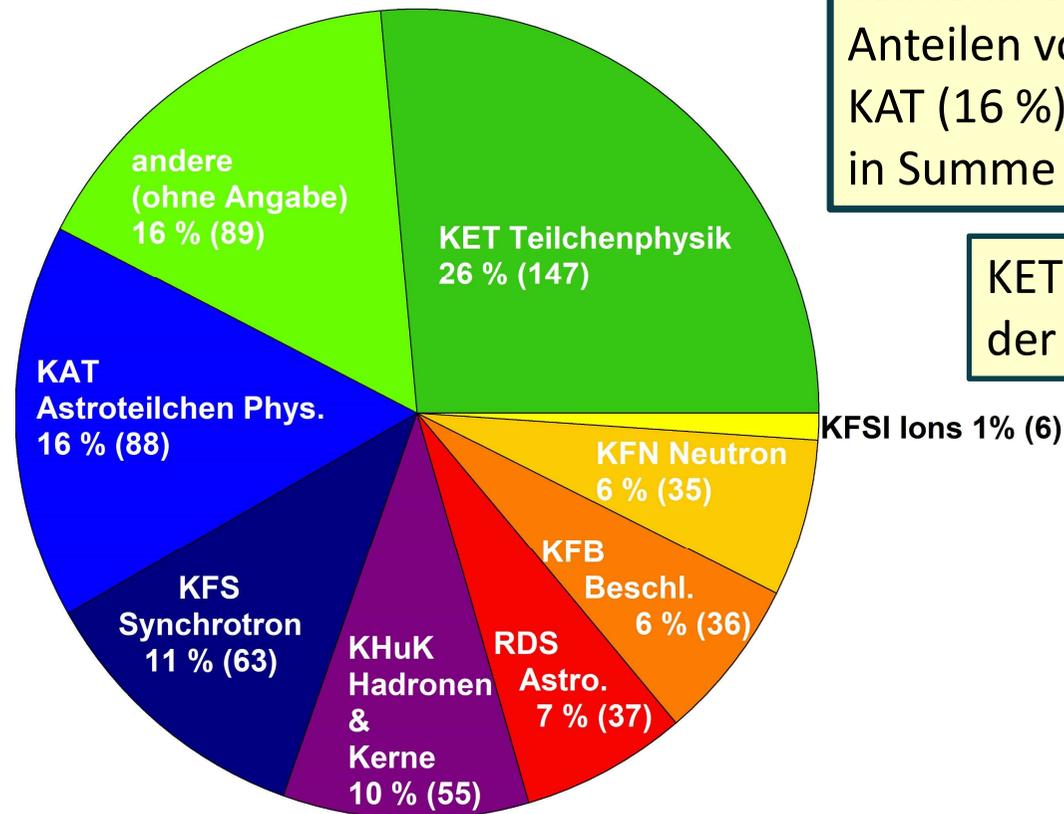
Nur zu geringen Anteilen:
Berlin, neue Bundesländer

⇒ Organisation von Workshops & Schulen in unterrepräsentierten Regionen in 2023 & 2024

⇒ Positive Wirkung eines „distributed office“ des Hubs?

Retrospektive: 3/2022 - 12/2023

Evaluation aller Veranstaltungen:



Teilnehmer zu großen Anteilen von KET (>25 %), KAT (16 %) & KFuK (10 %), in Summe > 50 %

KET + KAT + KFuK = 36 % der 8 ErUM-Community

KFS (11 %) & KFN (6 %) substantiell unterrepräsentiert

Zur Erinnerung:

KFS + KFN = 43 % aller ErUM-Wissenschaftler!

Retrospektive: 3/2022 - 12/2023



Σ 15 Events mit >520 Teilnehmenden

- 5 Deep learning schools
- 6 Expert Workshops
- 4 Train-the-Trainer Workshops
- 100% Weiterempfehlung



Teilnehmende vorwiegend aus dem Umland

- ⇒ In Zukunft: Organisation von Workshops, Schulen & Meetings in unterrepräsentierten Regionen
“neue Orte – neue Netzwerke”
- ⇒ ausreichend Reisemittel für bewilligte Projekte um Mobilität zu erzielen



Einige Komitees unterrepräsentiert

- ⇒ In Zukunft: Workshops zu stärker spezialisierten Themen
- ⇒ Bedarfsermittlung (siehe auch Talks von T. Kuhr, A. Schneidewind, P. Schnizer – große Bedarfe u.a. zu BDA, RDM, ...)

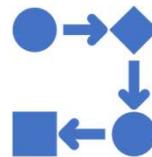
Verbreitung von Wissen stärken & junge Talente fördern



Data-Science Themen in den Universitätscurriculae

- Lehre von Digitalen Kompetenzen
- Dauerhafte Verankerung von Inhalten in den ErUM-bezogenen Studiengängen (z.B. zu Data Management, ELN, ...)
- Beginn bereits im Bachelorstudium

(siehe z.B. Physik Journal 21 (2022) 42)



Entwicklung von “tailored-learning” Konzepten

- Erarbeitung von Konzepten zur Wissenvermittlung unterschiedlicher Themen (z.B. Problem-based learning zu ausgewählten Themen und ermittelten Bedarfen)
- Veröffentlichung und Speicherung der Konzepte
- Nachhaltigkeitsproblematik



Qualifizierung des Nachwuchses

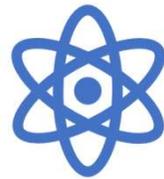
- Coaching Angebote
- Hilfe zur Selbsthilfe (z.B. Working out loud)
- Onboarding-Aktivitäten
- Aufbau von Community-Netzwerken
- Teilnahme an Workshops & Schulen
⇒ Finanzierung

Wahrnehmung von ErUM - Data



Einbeziehen der breiten Öffentlichkeit

- Podcast: 6 Folgen sind finanziert. Erste Folge in Arbeit
- Hannover-Messe
- Outreach-Konzepte in Anträge einarbeiten ⇒ Finanzierung notwendig



Wissensvermittlung

- Videos der Train-the-Trainer Workshops und Schulen (dauerhaft auf ErUM-Data-Hub abrufbar)
- Transfer auch in die Industrie
- Kontinuierliche Ermittlung aktueller Bedarfe
- ...



Verankerung von Knowledge Distribution in den Ausschreibungen

- Jedes geförderte Projekt sollte sich an Konzepten der Community beteiligen oder eigene Konzepte initiieren
- Finanzierung seitens des BMBF – Reisemittel für teilnehmende Wissenschaftlerinnen & Wissenschaftler

Zusammenfassung & Ausblick

- Existierende & erfolgreiche Aktivitäten weiterführen
(Deep-learning schools, TTT workshops, Experten-Workshops)
- Neue Formate und Themen etablieren:
 - Moderne Methoden
 - Programmierschulen
 - Developers Workshops
 - KI-erzeugter Programmcode
 - Nachhaltige Software
 - Transfer
 - ...
- Netzwerk-Aktivitäten unterstützen

Ihre Beiträge / Vorschläge sind willkommen & wichtig!

ERUM-DATA-HUB EVENT PROGRAMME 2024

- 26 FEB** **DEEP LEARNING SCHOOL**
- **1 MAR** **BASIC CONCEPTS** WIEHL
- 25 MAR** **DEEP LEARNING**
- **26 MAR** **TRAIN-THE-TRAINER** DRESDEN
- 21 MAY** **ADVANCED DEEP LEARNING**
- **24 MAY** **ACTIVE TRAINING COURSE** FILDERSTADT
- 19 AUG** **FAST AND EFFICIENT PYTHON**
- **22 AUG** **COMPUTING SCHOOL** AACHEN
- 26 AUG** **pyHEP.dev**
- **30 AUG** **DEVELOPER'S WORKSHOP** AACHEN

MORE TO BE ANNOUNCED SOON!

- > Conceptual Advances in Deep Learning
- > Software & Tooling Workshop

22 APR - 26 APR
HANNOVER MESSE

Registration & Contact:
www.erumdatahub.de
info@erumdatahub.de

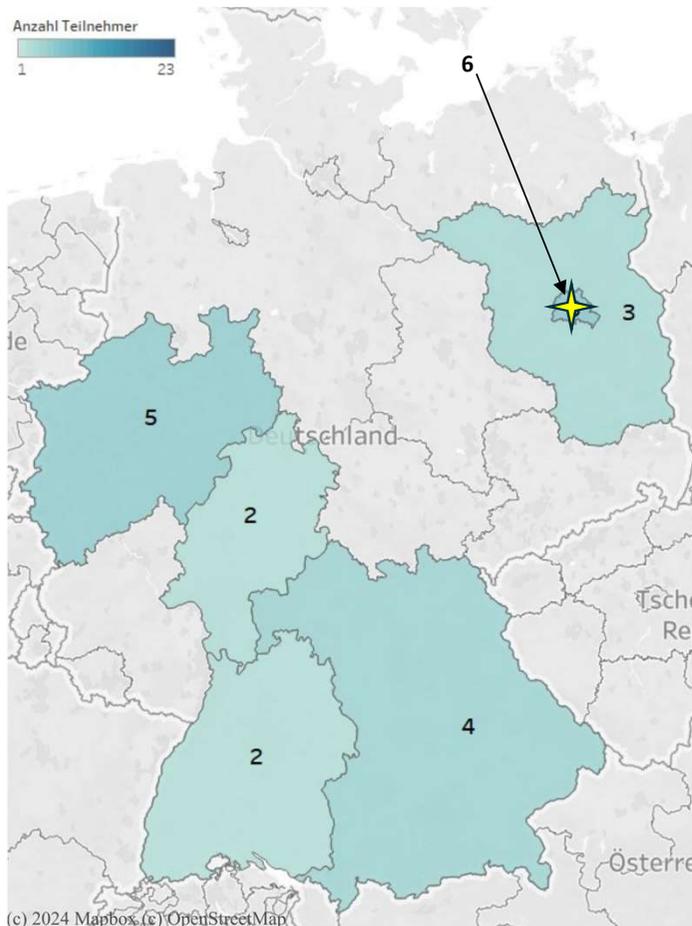


Follow us on

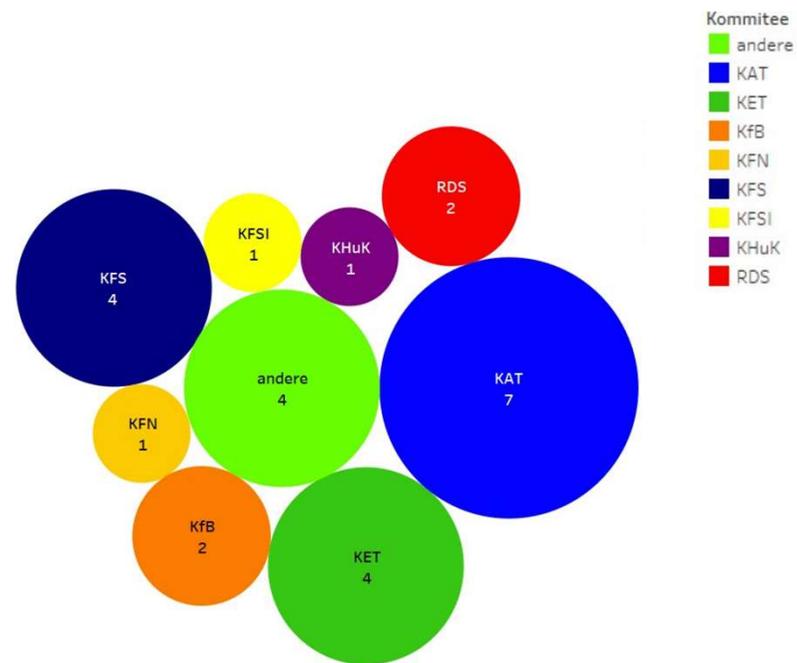




Expert Workshop Berlin: User Interface 5/2023



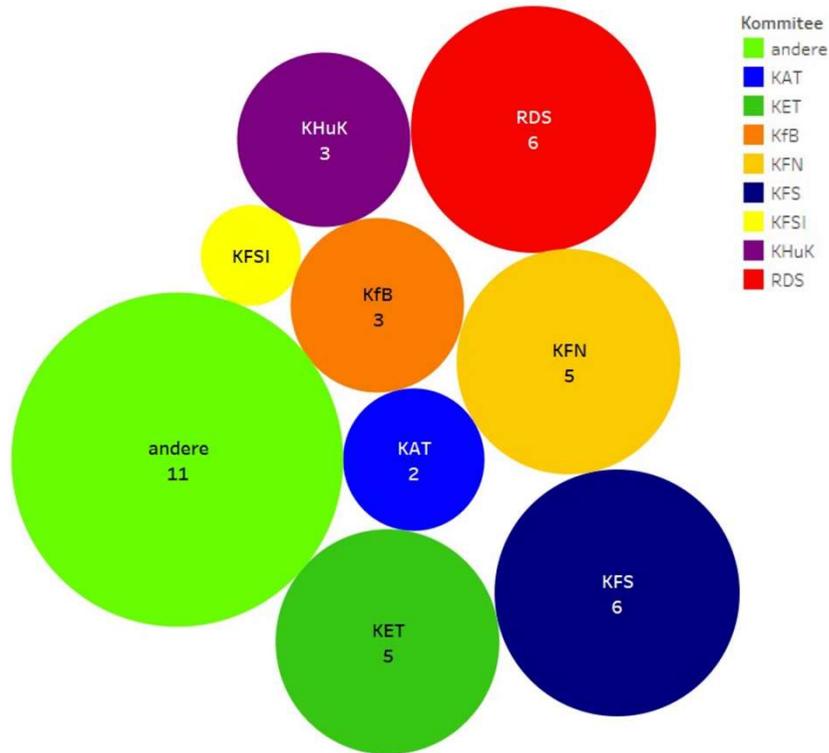
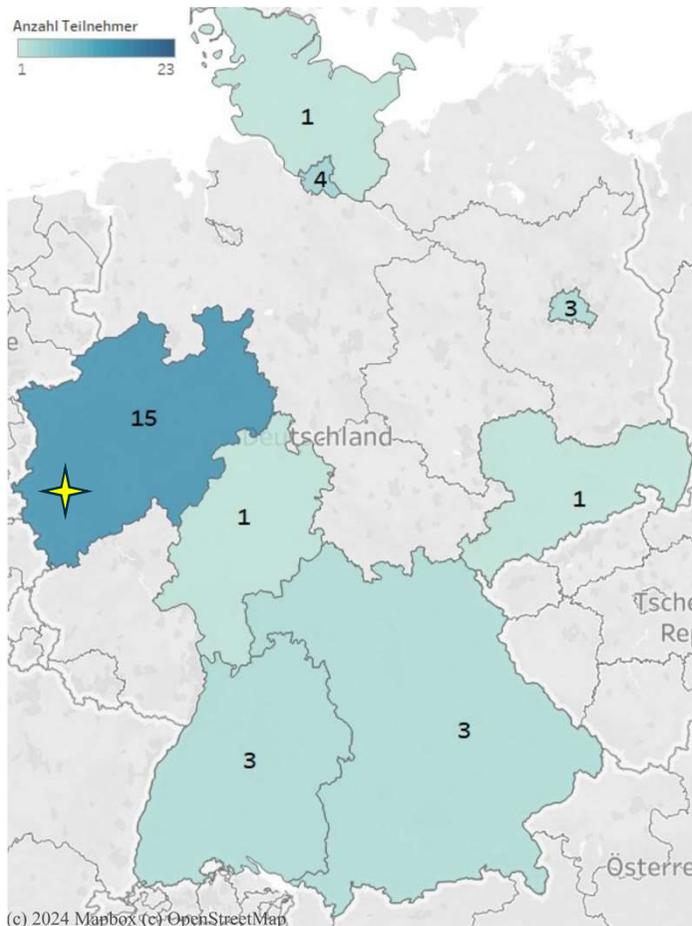
15



Σ 25 Teilnehmende

18.01.2024

Expert Workshop Bonn: Research Data Management 3/2023



Σ 32Teilnehmende

Retrospektive: 3/2022 - 12/2023



Σ 15 Events mit >520 Teilnehmenden

5 Deep-Learning-Schulen

6 Experten-Workshops

4 Train-the-Trainer Workshops

100% Weiterempfehlung



Detaillierte Evaluation der Veranstaltungen

Beispielhaft anhand von zwei Experten-Workshops:



Big Data Analytics - Connection Workshop

23.–24. Feb. 2023
DESY
Europe/Berlin Zeitzone

Geben Sie Ihren Suchbegriff ein

RDM - Expert Workshop on synergies between NFDI, ErUM-Data and Helmholtz-DMA

13.–14. März 2023
Max-Planck-Institut für Radioastronomie
Europe/Berlin Zeitzone

Geben Sie Ihren Suchbegriff ein

Next Generation Environment for Interoperable Data Analysis - Expert Workshop

3.–4. Mai 2023
HZB Berlin
Europe/Berlin Zeitzone

Geben Sie Ihren Suchbegriff ein

ErUM-Data Workshop on Inverse Problems

5.–6. Dez. 2023
Campus Garching
Europe/Berlin Zeitzone

Geben Sie Ihren Suchbegriff ein

Übersicht

- Tagesordnung
- Anmeldung
- Umfragen
- Payment
- Venue & Travel

Due to the difficult travel conditions, the workshop is also being broadcasted online. Join here: <https://rwth.zoom-x.de/j/67466745362>

This Workshop is hosted by the community organization [DIG-UM](#) with support from the BMBF-funded [ErUM-Data-Hub](#).

The goal of the workshop is to get an overview on inverse problems in the [ErUM-Data community](#), to discuss methods to solve them, and to foster exchange and cooperation between the ErUM-communities. In particular, the workshop is intended for early stage researchers.