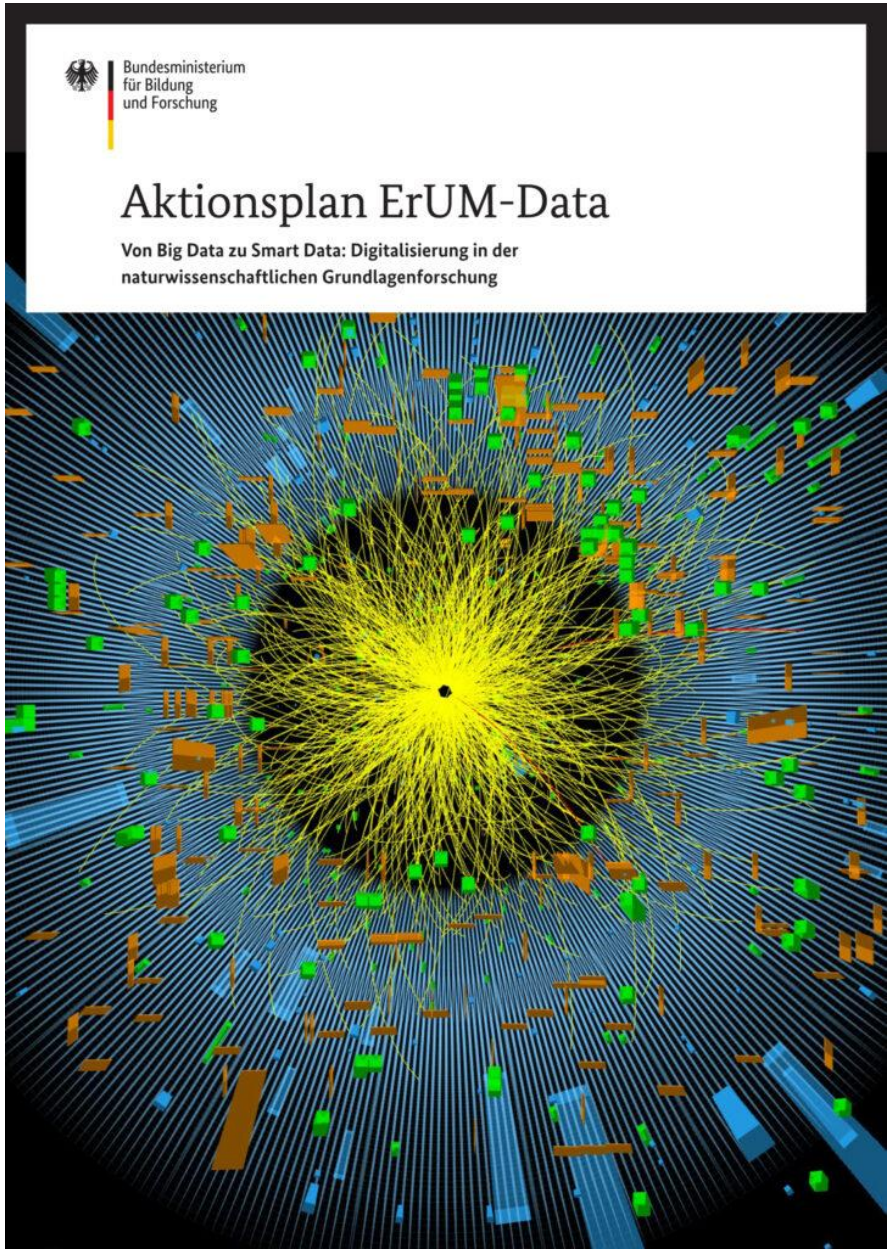
ErUM – DATA.
Preparation for new calls for funding
BMBF Prisma Trialog on Sustainability

Astrid Schneidewind, JCNS at MLZ – FZ Jülich



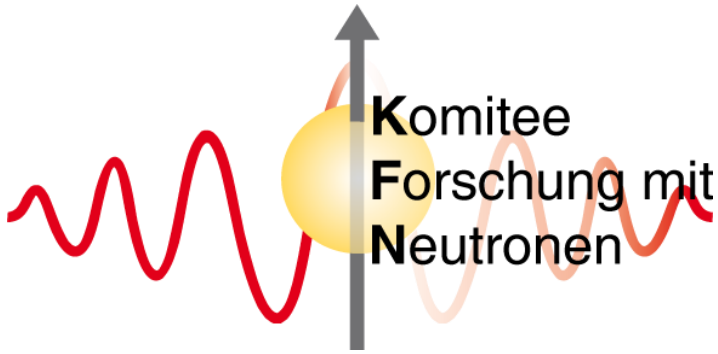
Challenges and Opportunities of Digital Transformation in Fundamental Research on Universe and Matter

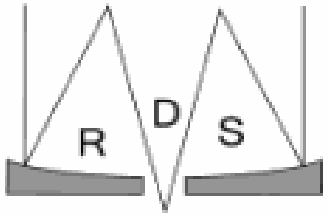
Recommendations of the ErUM Committees

[ErUM - Exploration of the Universe and Matter]

29 April 2019


KAT. Komitee für
Astro. Teilchen. Physik

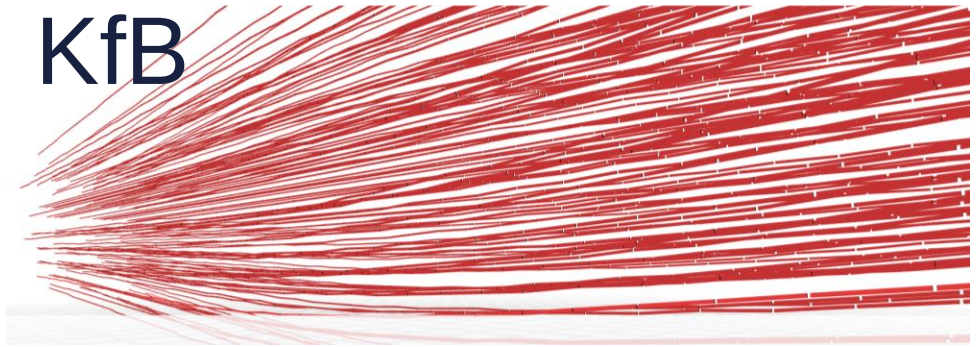

Komitee
Forschung mit
Neutronen


R
D
S


KET
Komitee für
Elementarteilchenphysik


Komitee
Forschung mit
Synchrotronstrahlung

KHuK


KfB


Komitee
Forschung mit nuklearen
Sonden und
Ionenstrahlen

ErUM Data is a BMBF funding scheme for these scientific communities, providing infrastructures for research on large scale facilities.

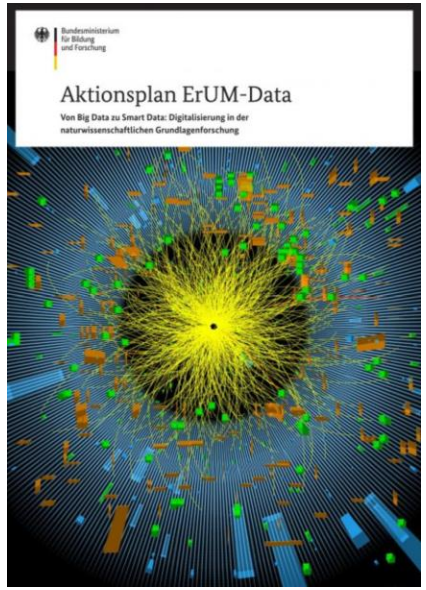
In contrast to NFDI the committees are requested by BMBF to cooperate on digitization in the interest of their overall community.

DIG-UM

Digital Transformation on Research of Universe and Matter

- ErUM Self-Organization for Digital Transformation, of the approx. 20,000 ErUM scientists
- Within the **self-organization**, the needs and requirements of the eight ErUM communities are combined, scientific strategies are developed and overview documents are written in different topic groups.
- Scientific strategies (Topic Groups)
- White papers, exchange, organize collaborations, at least support
- Founded 2-Nov-21 / 20-Jan-22

www.erumdatahub.de

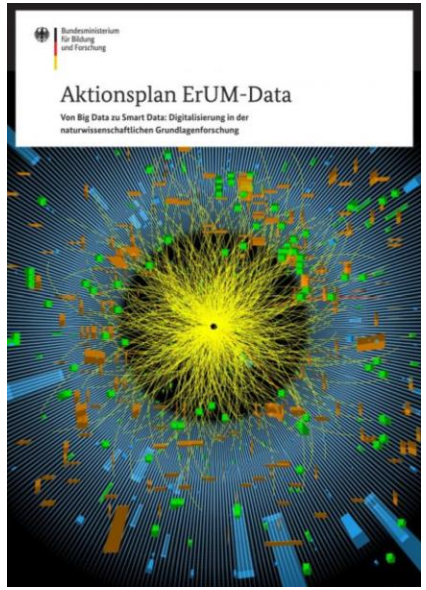


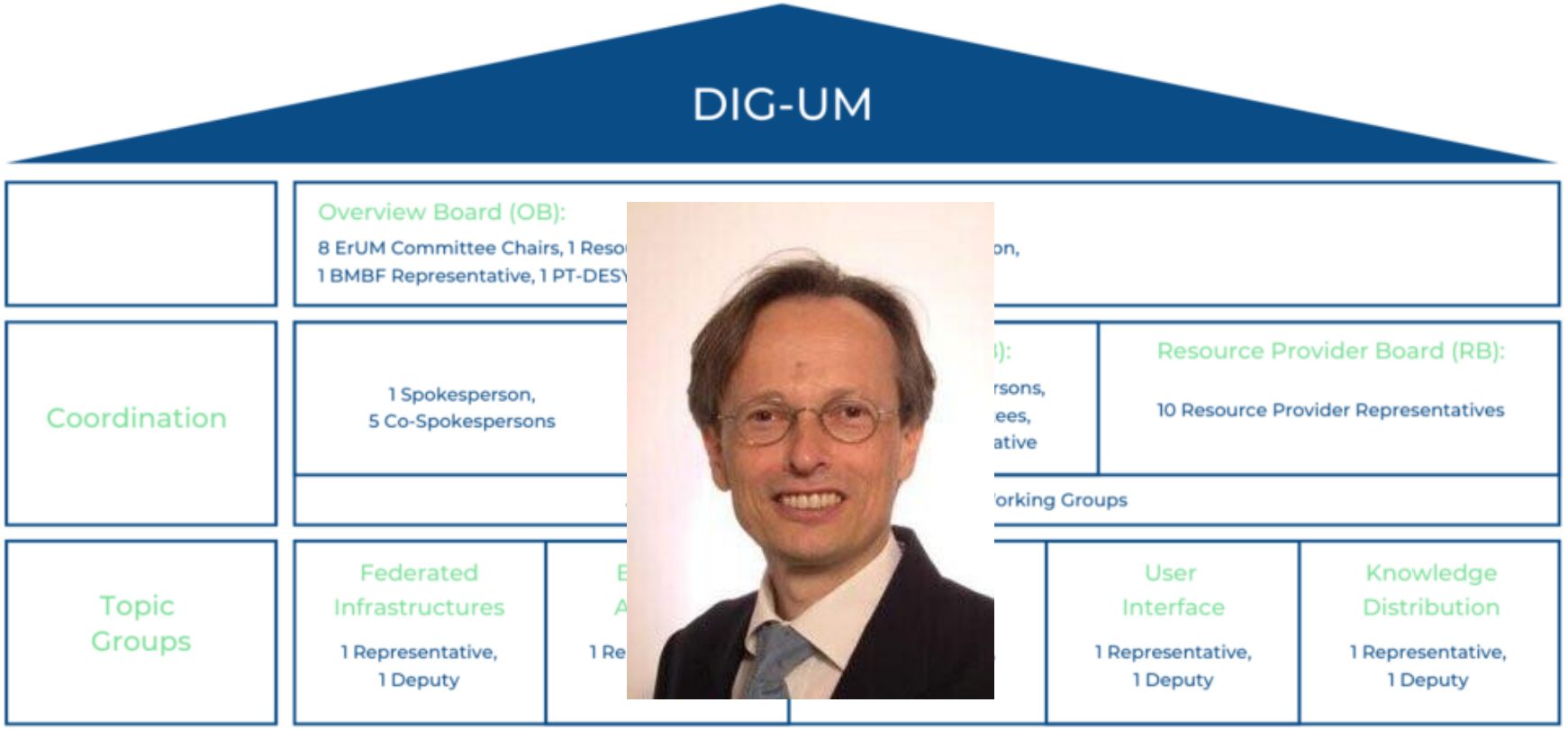
DIG-UM

	Overview Board (OB): 8 ErUM Committee Chairs, 1 Resource Provider Board Chair, 1 Spokesperson, 1 BMBF Representative, 1 PT-DESY Representative				
Coordination	1 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: 1 Representative, 1 Deputy	Big Data Analytics: 1 Representative, 1 Deputy	Research Data: 1 Representative, 1 Deputy	User Interface: 1 Representative, 1 Deputy	Knowledge Distribution: 1 Representative, 1 Deputy

DIG-UM unites the interests of the 8 research communities in the exploration of the universe and matter: Astroparticle Physics (KAT) Elementary Particles Physics (KET), Accelerator Physics (KfB), Research with Neutrons (KFN), Research with Synchrotron Radiation (KFS), Research with nuclear probes and ion beams (KFSI), Hadron and Nuclear Physics (KHuK) and the German Observatory Council (RDS).

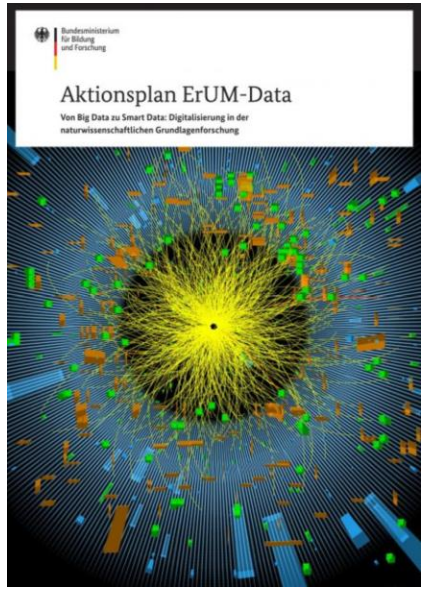
www.erumdatahub.de

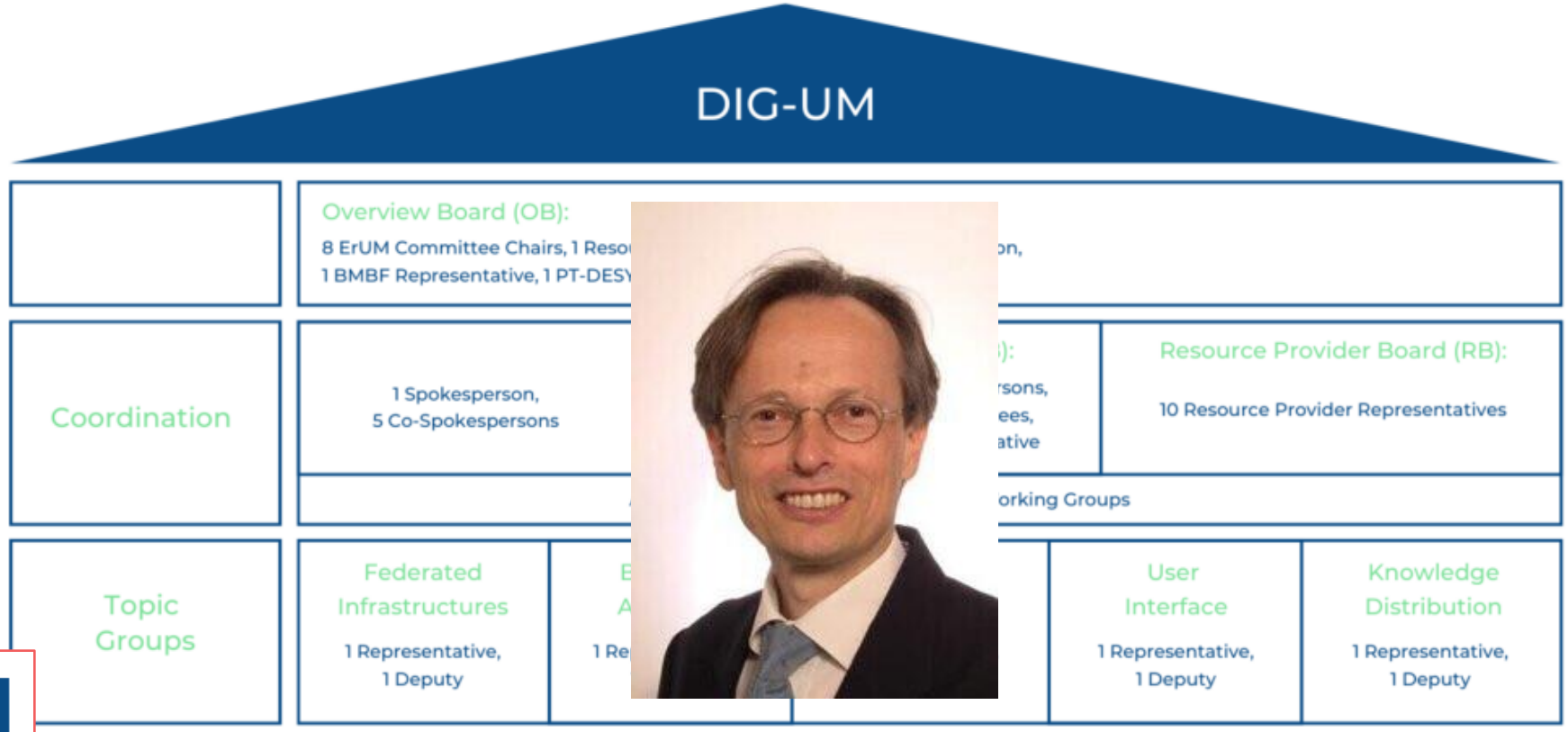




DIG-UM unites the interests of the 8 research communities in the exploration of the universe and matter: Astroparticle Physics (KAT) Elementary Particles Physics (KET), Accelerator Physics (KfB), Research with Neutrons (KFN), Research with Synchrotron Radiation (KFS), Research with nuclear probes and ion beams (KFSI), Hadron and Nuclear Physics (KHuK) and the German Observatory Council (RDS).

www.erumdatahub.de





ERUM DATA HUB

The activities of the ErUM-Data-Hub are divided into 4 work areas:

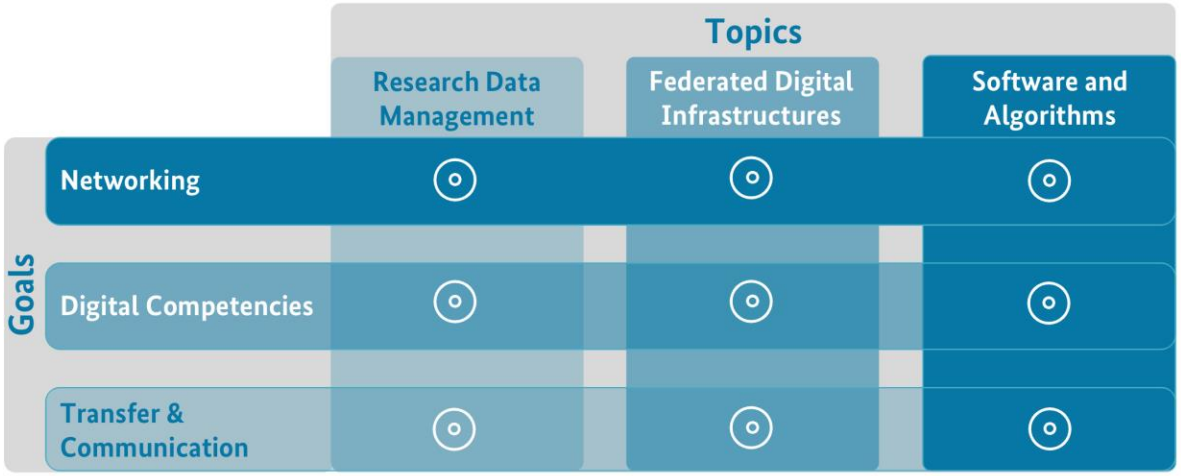
- Networking
- Exchange
- Digital competencies
- Communication

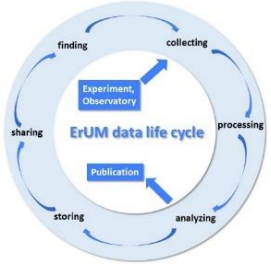
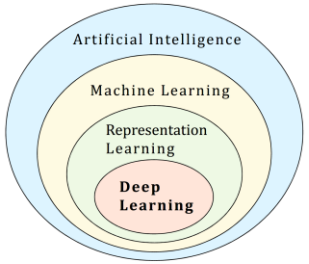
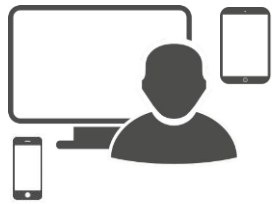
DIG-UM unites the interests of the 8 research communities in the exploration of the universe and matter: Astroparticle Physics (KAT) Elementary Particles Physics (KET), Accelerator Physics (KfB), Research with Neutrons (KFN), Research with Synchrotron Radiation (KFS), Research with nuclear probes and ion beams (KFSI), Hadron and Nuclear Physics (KHuK) and the German Observatory Council (RDS).

Matching community needs and funding schemes

Eckart Lilienthal, 20-Jan-22

ErUM-Data: Fields of action



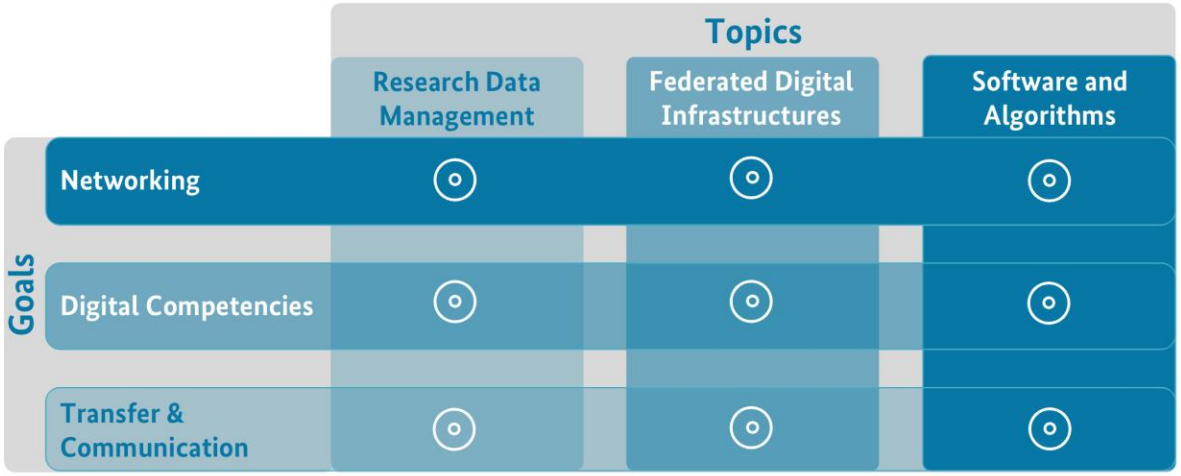
DIG-UM Topic Groups	Research Data	Federated Infrastructures	Big Data Analytics	User Interface	Knowledge Distribution
					

Matching community needs and funding schemes

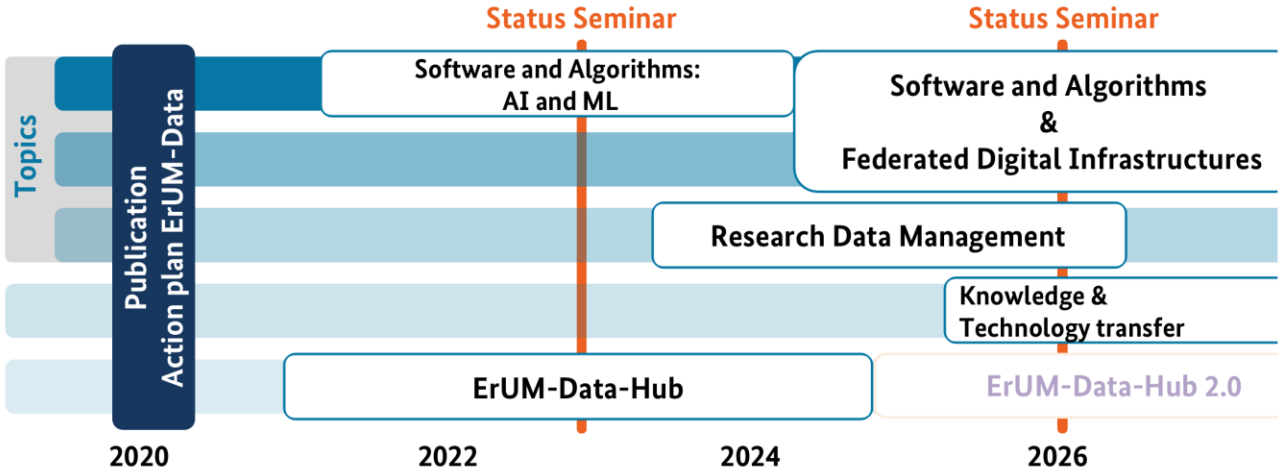
Eckart Lilienthal, 20-Jan-22

Bundesministerium für Bildung und Forschung

ErUM-Data: Fields of action

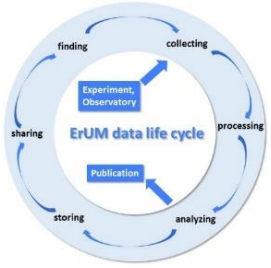


ErUM-Data: Timetable and implementation



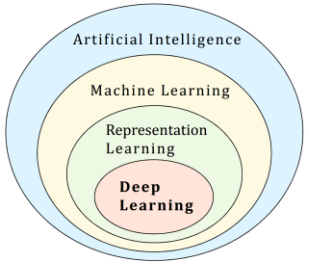
DIG-UM Topic Groups

Research Data

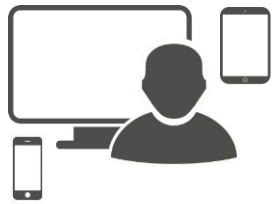


Federated Infrastructures

Big Data Analytics



User Interface



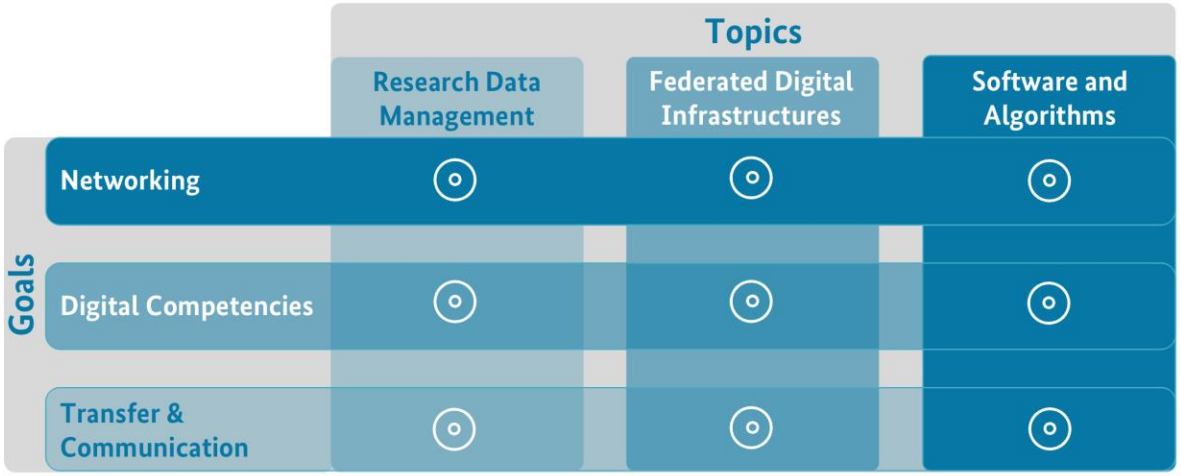
Knowledge Distribution

Matching community needs and funding schemes

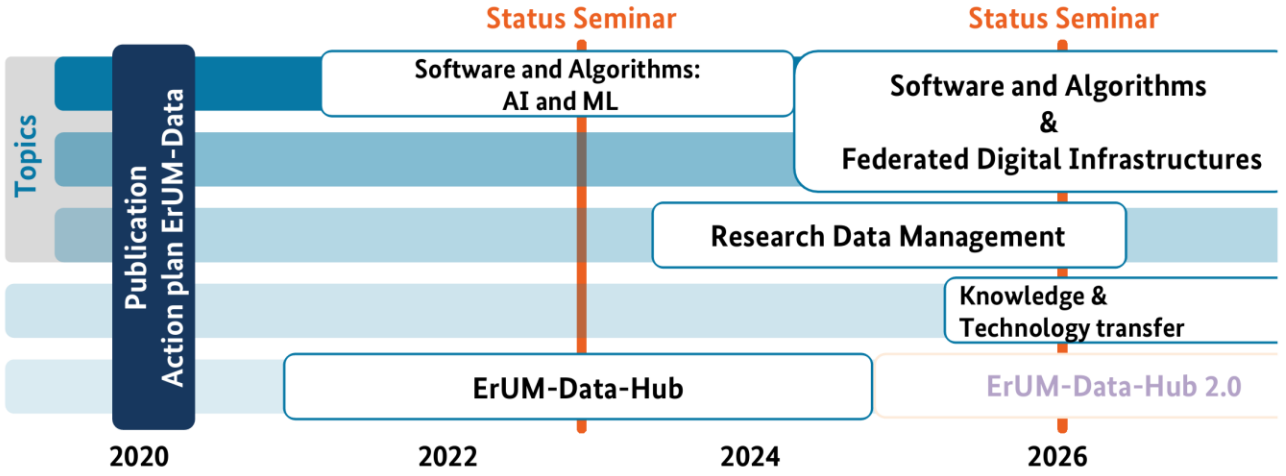
Eckart Lilienthal, 20-Jan-22

Bundesministerium für Bildung und Forschung

ErUM-Data: Fields of action

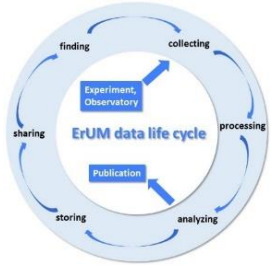


ErUM-Data: Timetable and implementation



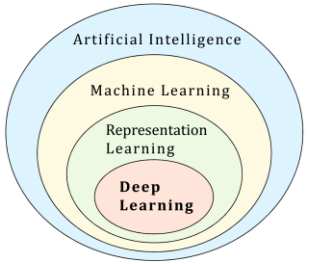
DIG-UM Topic Groups

Research Data

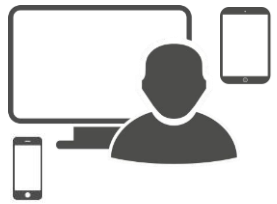


Federated Infrastructures

Big Data Analytics



User Interface



Knowledge Distribution

DIG-UM Topic Group Federated Infrastructures

Kilian Schwarz
Markus Demleitner

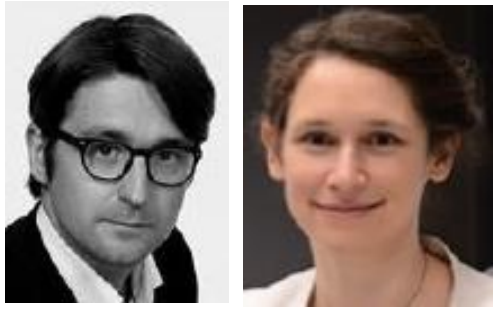
Mission

- Providing a distributed computing infrastructure:
  - in order to enable data taking, data processing and data archiving
  - including large data volumes and data of large diversity and
  - the required network backbone with sufficiently high bandwidth
- Computing resources need to support a larger number of users → need to be combined to federations. **easy and efficient way** for ErUM scientists to **access and use** these resources. The mid term target therefore needs to be to create an **ErUM Data Science Cloud** consisting of such a merger of German computing resources.

White paper identifying
– Current computing and storage infrastructures
– Already federated infrastructures
– Current issues which need to be addressed
– Infrastructures which are planned to be federated

available

DIG-UM Topic Group Big Data Analytics



Gregor Kasieczka
Andrea Thorn

erum-data-big-data-analytics.
lists@rwthaachen.de

Mission

- Connect researchers active e.g. in open ErUM-Data call on Software and Algorithms
- Topical cross-disciplinary workshops
- Education on key topics – both technical / scientific and legal / support

Projects of first call started summer 2023
Workshops to support and coordinate (next December 5th-6th)
Preparation of second call 2024

DIG-UM Topic Group Research Data Management


Astrid Schneidewind
Monica Valencia-Schneider

Mission

- Defining RDM as a topic to be funded
  - within the upcoming call and
  - within the environment of other funding schemes / projects as there are NFDI, Helmholtz incubator, DMA
  - Connect researchers active e.g. in open ErUM-Data call on Software and Algorithms
- Synergies
- Get researchers connected and activated

Synergy workshops connecting partners from NFDI, Helmholtz, DKZ…
Input for strategy discussion at BMBF January 23rd-24th 2024
Five workshops in February 2024 on local basis to inform the scientists and communities
Support cooperations (finding partners)

DIG-UM Topic Group User interfaces

Pierre Schnizer
Tim Ruhe

Goal and Challenges

- Develop science oriented web system / work bench
- Support modern scientific work flow

- Comfortable look and feel

Scientist: answering scientific questions.

- Technical support: comprehensive
- Technology: simplify scientific work
- Creativity: facilitate and support
- New ideas boost

No extra call in 2024 – Integrate topic in RDM call (and others?)

Web interface

Big Data Analytic Tools

Algorithms
Visualization
Machine Learning

Data

Experiment data
Metadata
Simulations

DIG-UM Topic Group Knowledge Distribution

Dirk Lützenkirchen-Hecht
Katrin Link

Topic	Target group	Participants	Date
Train-the-Trainer Deep Learning 1, Aachen	Lecturer	20	30.-31.3.22
Train-the-Trainer Deep Learning 2, Wuppertal	Lecturer	34	9.-10.6.22
Deep Learning Basic Concepts, Meinerzhagen	Beginners	57	8.-12.8.22
Conceptual Advances, Wiehl	Advanced	34	12.-16.9.22
Workshop NFDI & ErUM-Data, Frankfurt Airport	Experts	30	24.10.22
Deep Learning Advanced Concepts, Kloster Damme	Advanced	41	28.11.-1.12.22

Topic	Target group	Date
Deep Learning Basic Concepts	Beginners	February 23
Next Generation Environment for Interoperable Data Analysis	Experts	March 23
Train-the-Trainer Deep Learning Basic Concepts	Lecturer	March 23
Stand auf der Hannover Messe ErUM-Data-Hub, LHC-, Belle2-Büros	Public	17.-21. April 22
Train-the-Trainer Deep Learning Advanced Concepts	Lecturer	June 23
Deep Learning Basic Concepts	Beginners	August 23
Sustainability in the Digital Transformation	Experts	September 23
Deep Learning Advanced Concepts	Advanced	September 23

Deatiled program at the ErUM-Data-Hub website

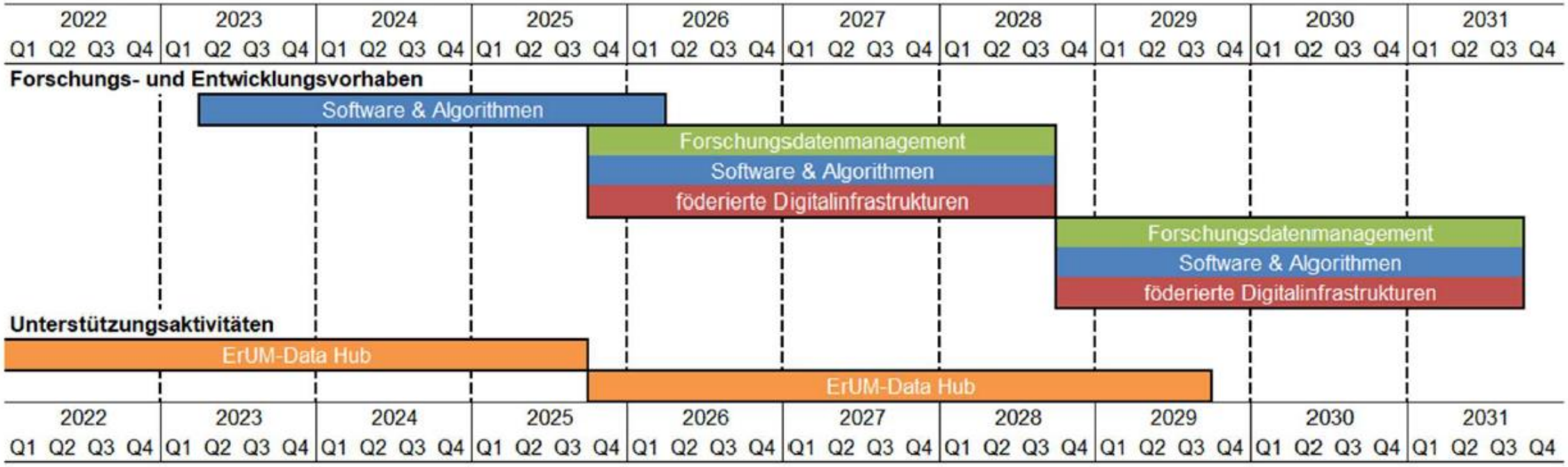
2023 / 2024 – change:
more local workshops
to attract more scientists

ErUM-Data: Preliminary schedule and implementation

- **Funding of the ErUM-Data Hub**
- **Calls and Funding of thematic R&D Projects**
- **„Software & Algorithms“ funding period 3/2023-2/2026**
- **„Federated Infrastructures“ pilot project FIDIUM (KET/KHuK) 6/2022 3y**

	No. joint projects (Verbünde)	No. of projects (Vorhaben)	(Requested) funding
Sketches (Skizzen)			74.5 Mio. €
Full proposals (Anträge)	33% of sketches		30.7 Mio. € 41% of sketches
funded	19% of sketches / 59% of full proposals	59% of full proposals	~17.5 Mio. € 23% of sketches 57% of full proposals

Software & Algorithms I



From: E.Lilienthal KAT Meeting Bad Honnef, 2 Dec 2022

Courtesy: A. Haungs

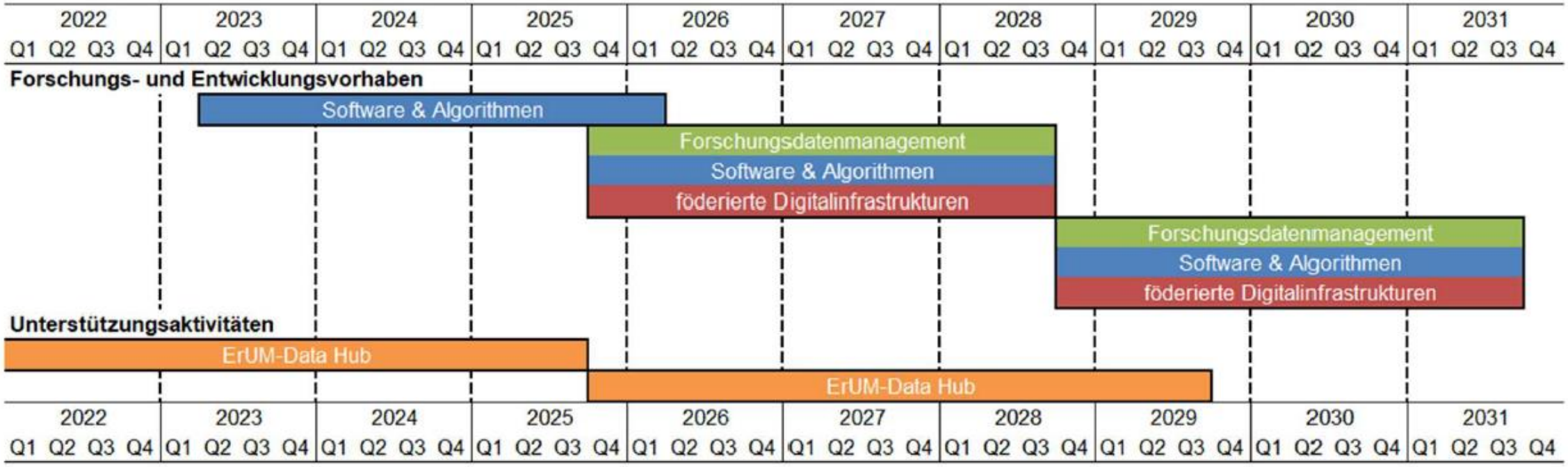
ErUM-Data: Preliminary schedule and implementation

- **Funding of the ErUM-Data Hub**
- **Calls and Funding of thematic R&D Projects**
- **„Software & Algorithms“ funding period 3/2023-2/2026**
- **„Federated Infrastructures“ pilot project FIDIUM (KET/KHuK) 6/2022 3y**

	No. joint projects (Verbünde)	No. of projects (Vorhaben)	(Requested funding)
Sketches (Skizzen)			74.5 Mio. €
Full proposals (Anträge)	33% of sketches		30.7 Mio. € 41% of sketches
Funded	19% of sketches / 59% of full proposals	59% of full proposals	~17.5 Mio. € 23% of sketches 57% of full proposals

2-step-procedure!!!

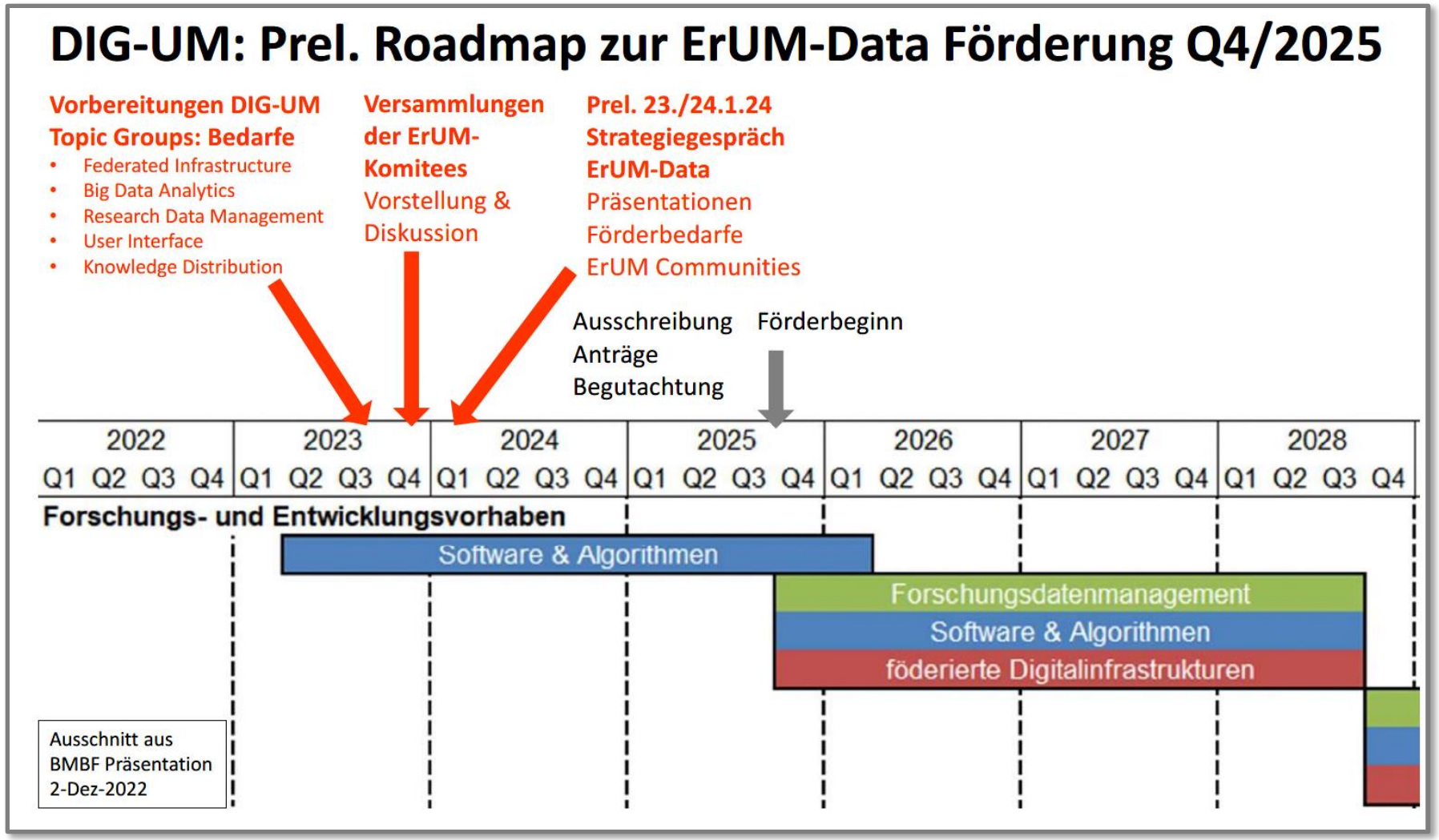
Software & Algorithms I



From: E.Lilienthal KAT Meeting Bad Honnef, 2 Dec 2022

Courtesy: A. Haungs

Preparation of strategy discussion in January



From: M. Erdmann, DIG-UM

- „Classical“ BMBF strategy discussion with the community for input to the call in Jan 2024
- Presentations of Topic Groups with input from Communities
- Invitation to the meeting also for PUNCH4NFDI and DAPHNE

Courtesy: A. Haungs

Prisma-Trialog des BMBF (ErUM): „Sustainability in research at large scale facilities: Resource efficiency & saving the future“

WG 1: Forschungsplanung und Organisation (Koordination: K. Haas)

WG 2: Forschungsförderung in ErUM (Koordination: F. Trinkel)

WG 3: Daten und Computing (Koordination: M. Erdmann, B. Murphy, M.Schumacher)

WG 4: Technologien an FIS (Koordination: E. Bründermann)

WG 5: Datenerhebung, Monitoring und Bilanzierung (Koordination: N.N.)

WG 6: Forschung für Nachhaltigkeit (Koordination: J.-D. Grunwaldt)

https://indico.desy.de/e/WGX_Trialog_Nachhaltigkeit

- **First Meeting in May 2023 with generation of WGs (tasks see next slide).**
- **Registration to WG‘s still possible….**
- **Second Trialog discussions Spring 2024**
- **➔ Decisions / Strategy for funding…?**

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

Ben Bruers[1], Marilyn Cruces[2], Markus Demleitner[3], Guenter Duckeck[4], Michael Düren[5], Niclas Eich[6], Torsten Enßlin[7], Johannes Erdmann[6], Martin Erdmann[6*], Peter Fackeldey[6], Christian Felder[8], Benjamin Fischer[6], Stefan Fröse[9], Stefan Funk[10], Martin Gasthuber[1], Andrew Grimshaw[11], Daniela Hadasch[9,12], Moritz Hannemann[8], Alexander Kappes[2], Raphael Kleinemühl[13], Oleksiy M. Kozlov[14], Thomas Kuhr[4], Michael Lupberger[15], Simon Neuhaus[13], Pardis Niknejadi[1], Judith Reindl[16], Daniel Schindler[17], Astrid Schneidewind[8], Frank Schreiber[18], Markus Schumacher[19], Kilian Schwarz[1], Achim Streit[20], R. Florian von Cube[20], Rodney Walker[4], Cyrus Walther[9], Sebastian Wozniewski[17], Kai Zhou[21]

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

Courtesy: A. Haungs

Thank you for your interest!

info@erumdatahub.de

a.schneidewind@fz-juelich.de

Prisma-Trialog des BMBF (ErUM):

"Sustainability in large-scale equipment research: efficiency & securing the future".

Schwerpunkte der Arbeitsgruppen

WG 1: Forschungsplanung und Organisation

- Vernetzung und Austausch (auch international) in der Fachcommunity und zwischen den Communities; Vergleich
- Kommunikation zum Thema Nachhaltigkeit mit der Politik, Öffentlichkeitsarbeit; Verbreitung von Informationen über Maßnahmen und Technologien
- Implementierung von Nachhaltigkeit im Arbeitsalltag:
  - Nachhaltigkeit als integralen Bestandteil eigener Forschung stärken
  - Engagement von Mitarbeitenden aus der Forschung & Entwicklung unterstützen, ggf. mit dafür reservierten Kapazitäten
  - Verhaltensänderungen ermutigen (Dienstreisen, Scientific Computing, etc)
  - Gemeinsame Nutzung von Aufbauten und Geräten anregen; end-of-life-Planung
  - Implementierung bottom-up statt nur top-down ermöglichen

WG 2: Forschungsförderung in ErUM

- Vorgaben der Rahmenbedingungen durch BMBF; z.B. für den Bereich Beschaffungen
- nachhaltige Forschungsförderung i.S.v. langfristiger Perspektive
- Weitere Fördermöglichkeiten für Nachhaltigkeitsmaßnahmen, v.a. mit Blick auf ErUM-spezifische Themen (großen Anlagen); vieles andere wird schon woanders gefördert
- Finanzierungsmechanismen
- Förderfähigkeit von Aktivitäten & Personal
- Prüfung von Nachhaltigkeitsanstrengungen anhand konkreter Kriterien
- Spagat zwischen Exzellenz und Nachhaltigkeit: Fehlanreize bei Förderung vermeiden

WG 3: Daten/Computing

- Nachhaltigeres HPC & HTC
- Datenmanagement und Datenanalyse inkl. ML/KI
- Speicherformate
- Technische Aspekte (z.B. Abwärmenutzung, Kühlung, ...)
- Forschungskultur: Umdenken Richtung Sparsamkeit und Effizienzgewinne; Verankern von Nachhaltigkeit im Forschungshandeln
- Querschnittsthema mit Bezug zu
  - Open Data/Open Software
  - Reproduzierbarkeit von Forschungsergebnissen

WG 4: Technologien an FIS

- Innovationen in Technologien von Komponenten zum System (Beschleuniger, Magnettechnik, Energy Recovery LINACs etc.)
- Bau von Demonstratoren nötig – Förderung sinnvoll
- Transferpotentiale für Nutzung außerhalb der Forschung

WG 5: Datenerhebung, Monitoring und Bilanzierung

- Dopplungen vermeiden bei Monitoring und Berichten; Verpflichtungen und Vorgaben müssen besser abgestimmt werden, um Aufwand überschaubar zu halten
- Bilanzierung von Lieferketten
- Reduktion von Emissionen bei Beschaffungen
- Referenzsysteme für internationale Vergleichbarkeit; z.B. ERIC-Verordnung der EU-KOM: einheitliche Bilanzierung der Umweltwirkungen von FIS; (Umwelt-)Wirkungen zukünftiger FIS müssen von Beginn an mitbetrachtet werden

WG 6: Forschung für Nachhaltigkeit

- Nutzung von Photonen-, Neutronen- und Ionenquellen für Forschung für Nachhaltigkeit
- Kommunikation der Beiträge von Grundlagenforschung an FIS in der Öffentlichkeit

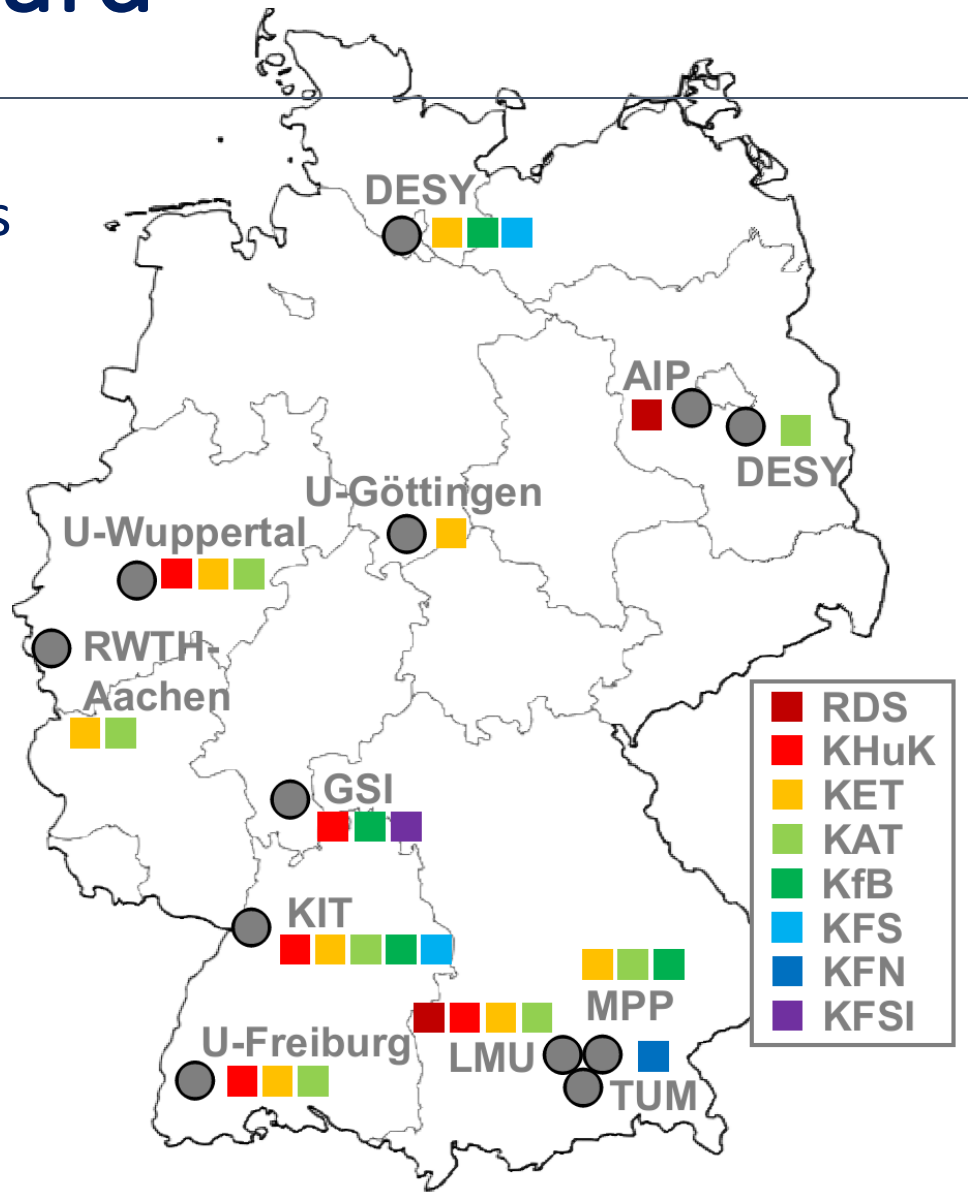
Courtesy: A. Haungs

DIG-UM **Resource Provider Board**

Achim Streit

Current resources for ErUM communities

In total
>280.00 CPU-cores
>700 GPUs
>260 PB disk and
>100PB tape storage



DIG-UM Topic Group Research Data Management


Astrid Schneidewind
Monica Valencia-Schneider

Topics of interest

Data Curation

- Curation vs. selection
- Scalability, extensibility
- Where, when, who?

Data storage

- Short term vs. long term
- Archiving: styles and techniques

In-situ analysis

- Reproducibility
- Version control

Technical implementation

- Storage and data management systems
- Computing architectures for RDM

Licencing, Embargos

- Publishing data
- What is open
- Who can use/re-use data

Open data

- Who is responsible for sharing the data ?
- Data analysis codes
- Who is responsible to make them FAIR?

EU requests

- Certification of repositories
- Legal aspects