A detailed wireframe model of a particle accelerator, likely the FAIR facility. The model shows a complex arrangement of circular and linear sections, with a large, prominent ring structure in the foreground. The background shows more intricate structures, possibly representing different stages or components of the accelerator.

# **Digitalisierung in BMBF Rahmenprogramm Erforschung von Universum und Materie (ErUM) - Participation of KHuK and foundation of a computing working group**

Kilian Schwarz

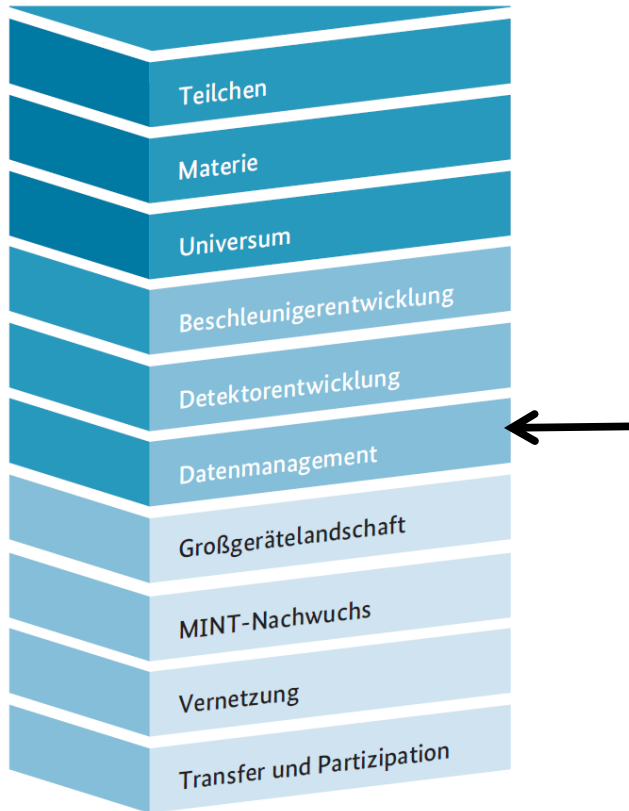
GSI

# Outline

- what is ErUM-Data
- ErUM-Data pilot project
- KHuK-Data group
- Computing in ErUM – BMBF workshop in Bonn
- NFDI
- related computing infrastructure projects

# Erforschung von Universum und Materie – ErUM

## Rahmenprogramm des Bundesministeriums für Bildung und Forschung



Naturwissenschaftliche Grundlagenforschung ist zentrales Einsatzfeld neuer digitaler Methoden und Techniken. Sie ist maßgeblicher Treiber für die weitere digitale Entwicklung. Steigendem Berechnungsaufwand und komplexem Datenmanagement wird mit standortübergreifenden Arbeitstechniken und der Beseitigung technologischer Flaschenhalse begegnet. Offener Zugang und langfristige Datenhaltung müssen dabei auch in Zukunft die Anforderungen und Spezifika der verschiedenen Forschungsinfrastrukturen berücksichtigen. Der wissenschaftliche Nachwuchs baut im Bereich Datenmanagement eine einzigartige Expertise auf. In Zukunft können auf Basis des Know-hows aus der Grundlagenforschung neue Dienstleistungen und ganzheitliche Lösungen entstehen.

- Particle physics, hadron and nuclear physics, and astro particle physics face similar huge technological challenges
  - Vast increase of data volumes
  - Evolution of technologies
- Expertise in key areas available at several German institutes
- Bundling of expertise in collaborative research project with partners from particle, hadron and nuclear, and astro particle physics **to develop experiment overarching solutions**
- Pilot project R&D Computing funded by BMBF with ~12 FTE for three years starting October 2018

# Project Partners



JOHANNES GUTENBERG  
UNIVERSITÄT MAINZ

Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG



JÜLICH  
FORSCHUNGSZENTRUM



# Subject Area A

› Developments for the provision of technologies for the use of heterogeneous computing resources

Item 1  
of  
strategy  
document

<p><b>A1) Werkzeuge zur Einbindung</b></p> <ul style="list-style-type: none"> <li>• Scheduling von Cloud-Jobs</li> <li>• Container-Technologien</li> <li>• Checkpointing</li> <li>• Zugang zu Experiment-Datenbanken</li> </ul>	<p><b>A2) Effiziente Nutzung</b></p> <ul style="list-style-type: none"> <li>• Transiente Datencaches</li> <li>• Transparenter Zugriff auf verteilte Daten</li> </ul>
<p><b>A3) Workflow-Steuerung</b></p> <ul style="list-style-type: none"> <li>• Identifikation und Steuerung</li> <li>• In-Pilot Job-Monitoring</li> <li>• Accounting</li> <li>• Optimierung durch Data-Mining</li> </ul>	

# Subject Area B

## > Application and test of virtualized software components in the environment of heterogeneous computing resources

Item 1  
of  
strategy  
document

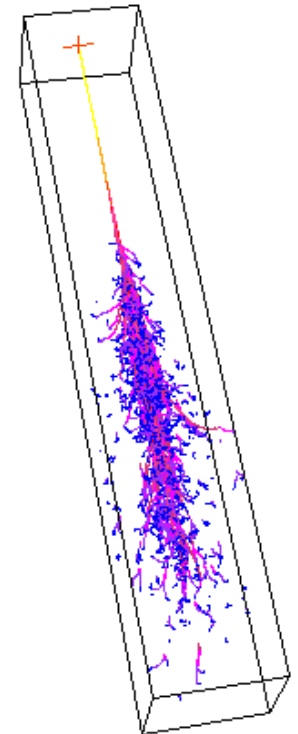
<p><b>B1) Tests der Technologiekomponenten</b></p> <p>Implementierung und Tests auf verschiedenen Plattformen von</p> <ul style="list-style-type: none"> <li>• Speicher- und Cachinglösungen und</li> <li>• virtualisierter Dienste (Datenbanken, Monitoring, Accounting).</li> </ul>	<p><b>B2) Job- und Ressourcenmanagement</b></p> <p>Jobverteilung und Überwachung in der Umgebung heterogener Computingressourcen unter Einbeziehung von Containervirtualisierung.</p>
<p><b>B3) Virtualisierung von Nutzerjobs</b></p> <ul style="list-style-type: none"> <li>• Erfassung der Anforderungen,</li> <li>• Bestimmung und Erzeugung der Laufzeitumgebung,</li> <li>• Erstellung des Containers und von Metadaten und</li> <li>• Checkpointing von Containervirtualisierung.</li> </ul>	<p><b>B4) Kombinierte Tests</b></p> <p>Testen von Gesamtsystemen (Speicher, Dienste, Ressourcenmanagement) auf verschiedenen Plattformen in Bezug auf</p> <ul style="list-style-type: none"> <li>• Installations- und Wartungsaufwand,</li> <li>• Performance,</li> <li>• Skalierbarkeit und</li> <li>• Robustheit.</li> </ul>

# Subject Area C

➤ Deep Learning, Gain of knowledge by substantiated data-driven methods

Item 2 of strategy paper

<p><b>C1) Sensornahe Verarbeitung von Daten</b></p> <ul style="list-style-type: none"> <li>• Signalfilter, Rauschunterdrückung</li> <li>• Verarbeitung von zeitabhängigen Signalen</li> </ul>	<p><b>C2) Objektrekonstruktion</b></p> <ul style="list-style-type: none"> <li>• Spur- und Clusterrekonstruktion, Jetbildung, Ereignisrekonstruktion</li> <li>• Fragestellungen für Anordnung, Reihenfolge, Zuordnungen von Daten</li> <li>• Optimierungen zur Extraktion kleiner Signale bei großem Untergrund</li> </ul>
<p><b>C3) Netzwerkbeschleunigte Simulationen</b></p> <ul style="list-style-type: none"> <li>• Generative adversarial networks, Anpassung von Simulationen an Datenverteilungen</li> <li>• Evaluationsverfahren für die Qualität der Netzwerksimulationen</li> </ul>	<p><b>C4) Qualität von Netzwerkvorhersagen</b></p> <ul style="list-style-type: none"> <li>• Reduzierung experimenteller systematischer Unsicherheiten</li> <li>• Spezielle Lernstrategien</li> <li>• Vorhersagenrelevante Information</li> <li>• Unsicherheiten von Vorhersagen</li> </ul>

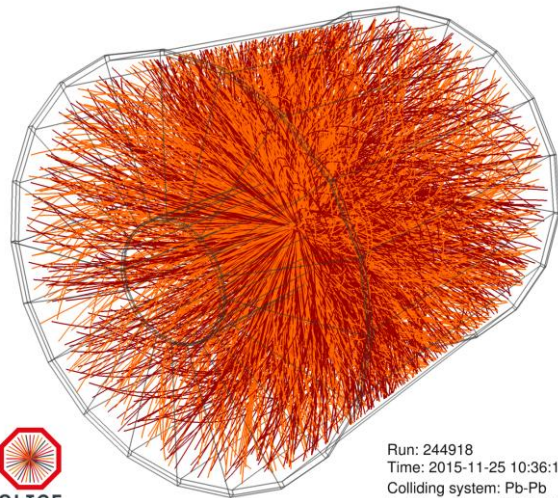




# Subject Area D

## Event reconstruction: Cost- and energy-efficient use of computing resources

Item 2  
of  
strategy  
document

<p><b>D1) Spurfindung</b></p> <ul style="list-style-type: none"> <li>• alternative Algorithmen, z.B. zellulärer Automat</li> <li>• alternative Architekturen, z.B. GPUs</li> </ul>	<p><b>D2) Parameterbestimmung</b></p> <ul style="list-style-type: none"> <li>• Verknüpfung GenFit2-ACTS</li> </ul>
<p><b>D3) Neutrinoexperimente</b></p> <ul style="list-style-type: none"> <li>• dünnbesetzte Detektorinstrumentierung</li> <li>• variable Signalzeit als kritische Information</li> </ul>	



Run: 244918  
Time: 2015-11-25 10:36:18  
Colliding system: Pb-Pb  
Collision energy: 5.02 TeV

# Overview Work Packages

Institute	Experiments	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3
Aachen Erlangen	CMS, Auger, $\nu$ -Exp. CTA, IceCube								X	X		X	X		X
Frankfurt	ALICE, CBM	X	X	X	X			X					X		
Freiburg	ATLAS	X	X	X		X		X							
Hamburg	CMS									X		X			
Karlsruhe	CMS, Belle II, Auger	X	X	X							X			X	
Mainz	ATLAS								X	X					
München	ATLAS, Belle II		X	X	X		X	X			X				
Wuppertal	ATLAS	X		X	X		X	X							
<i>Associated</i>															
CERN	ATLAS							X		X			X	X	
DESY	(ILC in WP D2)		X					X						X	
GridKa								X							
GSI		X	X					X							
Jülich	Panda													X	
Münster	ALICE, CBM									X	X				

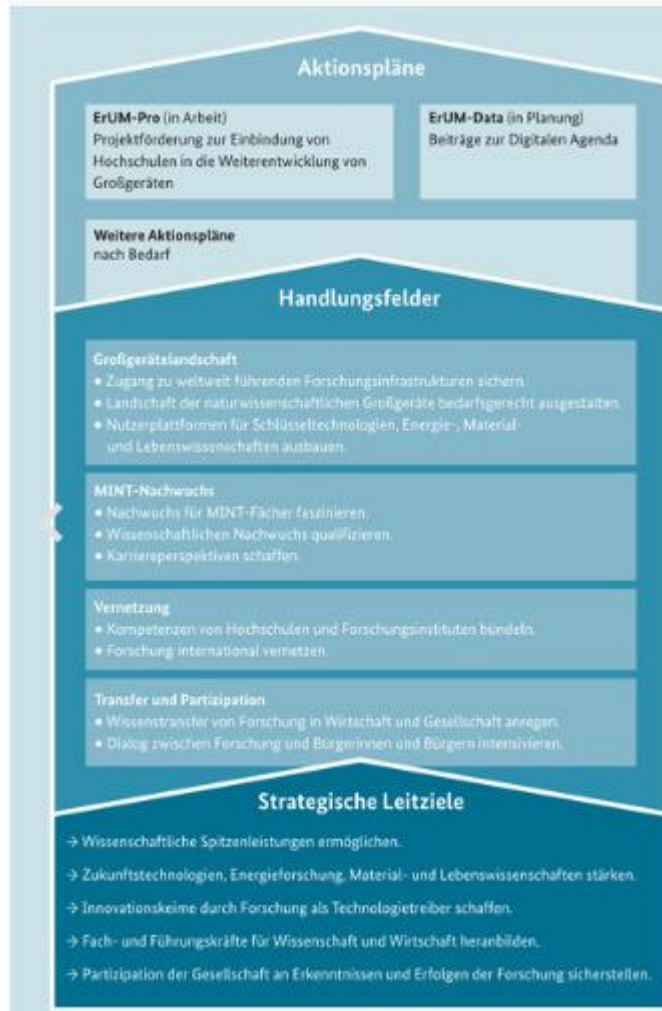
# Founding of KHuK Computing Group

## KHuK data, founded September 13, 2018



- in order to be able to address KHuK computing issues in a competent way
- in order to have a competent contact partner for computing issues coming from outside KHuK (BMBF, other committees (KAT, KET, ...))
- members:
  - Volker Friese (GSI)
  - Frithjof Karsch (Uni Bielefeld/BNL)
  - Ivan Kisel (FIAS/Uni Frankfurt)
  - Thorsten Kollegger (GSI/FAIR)
  - Marcel Kunze (Computing Centre Uni Heidelberg)
  - Frank Maas (HIM/GSI/Uni Mainz)
  - Peter Malzacher (GSI)
  - Silvia Masciocchi (Uni Heidelberg/GSI)
  - Walter Müller (GSI)
  - Kilian Schwarz (GSI)
  - Tobias Stockmanns (IKP Jülich)
  - Andrey Surzhykov (TU/PTB Braunschweig)
  - OPEN FOR MORE COMPUTING INTERESTED ACTIVE PARTICIPANTS
- KET already has an active computing panel, which helps KET to deal fast and efficiently with computing related topics and decisions are embedded in the community

- within the ErUM program the BMBF has put in place initiatives to discuss various aspects related to "Data": big data, storage, transport, network, infrastructure, analysis methods, fast algorithms, ....
- The BMBF organised a workshop to discuss these topics among the physics communities KAT, KET, KHuK, KfB, KFN, KFS, RDS, which happened on October 4-5 in Bonn.
- KHuK data was involved in the preparation of the workshop. KHuK data members participated and provided valuable input to the 3 working groups
  - Federated Infrastructures
  - Big Data Analytics
  - Research Data Management



- Outcome of the BMBF workshop will be a strategy document which will contain key points of the computing needs of all communities. The document will be communicated back to the communities for comments
- The content of this document will be the basis of an upcoming funding procedure for computing in ErUM
- **Be aware of things to come and be prepared to act fast !!!**
- applications should be done in consortia. Criteria for success are not only scientific excellence but also collaborative work and enlarging user communities

## **Bund-Länder-Vereinbarung zu Aufbau und Förderung einer Nationalen Forschungsdateninfrastruktur (NFDI) vom 26. November 2018**

**Bund und Länder stellen ... im Zeitraum  
von 2019 bis 2028 bis zu 90 Mio Euro ...  
für die Projektförderung der NFDI zur  
Verfügung**

KHuK, KET and KAT working on a common  
project proposal to connect their communities to  
the national research data infrastructure NFDI

small working group formed to  
work on this, KHuK Data as main  
contact point. First outcome will  
be a joint position paper.

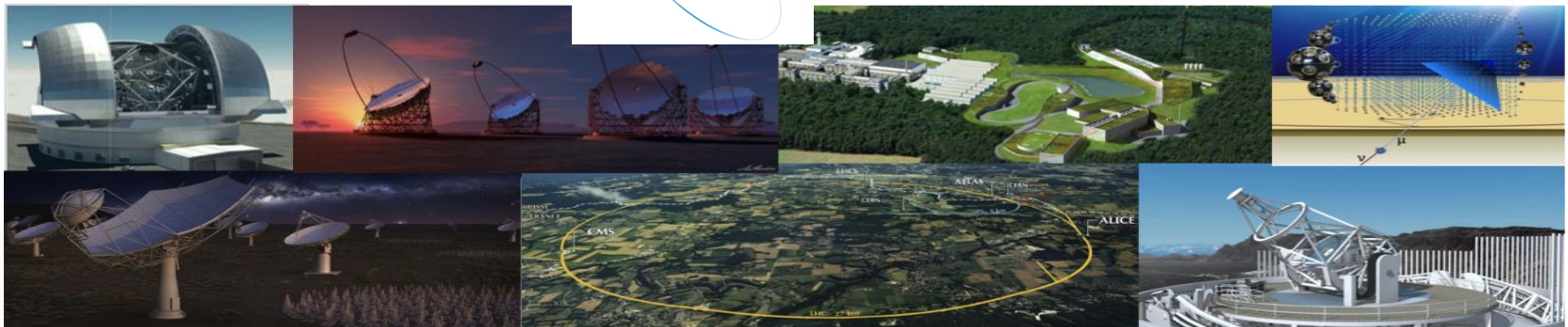
# KHuK already participates in similar computing infrastructure related projects:



Eine europäische Wissenschaftscloud verwirklichen, den universellen Zugriff auf Forschungsdaten über eine einzige Online-Plattform ermöglichen – das ist das Ziel der Initiative European Open Science Cloud (**EOSC**), die von den Mitgliedsstaaten der EU ins Leben gerufen wurde. Eine Förderzusage erhielt nun das EU-Projekt **ESCAPE** (European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures). Auch FAIR und GSI sind an diesem umfassenden Vorhaben entscheidend beteiligt.

## Helmholtz-Datenföderation (HDF)

Die Helmholtz-Datenföderation (HDF) ist eine strategische Initiative der Helmholtz-Gemeinschaft, die sich einer der großen Herausforderungen des nächsten Jahrzehnts annimmt: Die Bewältigung der Datenflut in der Wissenschaft, insbesondere aus den großen Forschungsinfrastrukturen der Helmholtz-Zentren.







# What is ErUM

- Richtlinie zur Förderung von ausgewählten Schwerpunkten der naturwissenschaftlichen Grundlagenforschung auf dem Gebiet „Erforschung der Materie an Großgeräten“ innerhalb des Rahmenprogramms „Erforschung von Universum und Materie – ErUM“; Bundesanzeiger vom 20.09.2018
- Daher ist Zweck dieser Maßnahme,
  - die überregionale Zusammenarbeit von Hochschulgruppen im Verbund mit den Großgeräten zu stärken sowie
  - die qualifizierte Nutzung der Großgeräte durch die Gewinnung und Ausbildung wissenschaftlichen Nachwuchses sicherzustellen