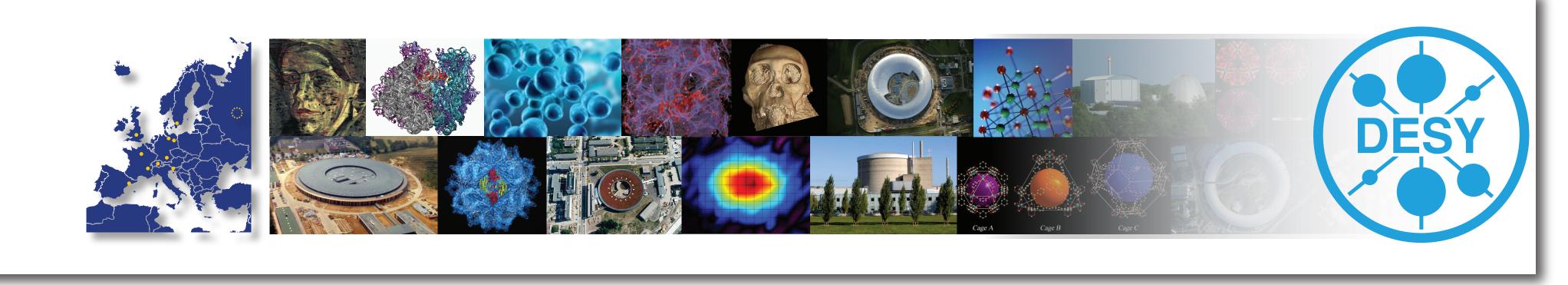
## Management Wissenschaftlicher Daten



Daten sind wertvoll! Der Verlust von Daten kann ein schmerzhaftes und kostspieliges Problem sein.

Das Zeitalter von Wechselplatten in der wissenschaftlichen Datenhaltung ist vorbei. Die Daten-Infrastruktur am DESY bietet ein sicheres Archiv für wissenschaftliche Daten, schützt vor Datenverlusten und ist von überall im Netz zugreifbar.

DESY bietet leistungsfähige und einfach zugängliche Rechner-Infrastrukturen für die Analyse der wissenschaftlichen Daten.

Zusammen mit anderen Photonen und Neutronenquellen in Europa bemüht sich DESY, eine gemeinsame umfassende Struktur für das Datenmanagement aufzubauen.

## Suchen und Finden



Um den wissenschaftlichen Nutzern die Suche in den großen Datenbeständen zu erleichtern, richten

wir einen föderierten Datenkatalog ein. Damit lassen sich die Datenarchive aller europäischen Synchrotron-

Forschungszentren durchsuchen, und benötigte Daten werden übers Netz zu den Analyserechnern geschickt.

Dafür muss der Datenkatalog alle Nutzer und deren Berechtigungen kennen.

entsprechendes System Ein bietet Nutzern weltweit eindeutige elektronische Ausweise, die an allen teilnehmenden Instituten gültig/

sind.

Managen

Für das Datenmanage-

ment betreibt DESY ein

Web-basiertes System, das

Langzeitarchivierung

die Rohdaten in ein für die

010110 110011 101000

Management

Computing

Catalog

**Education** 

Experiment Measurement Simulation

kann. Dort können Datenbestände

gee-

laden werden.

ignetes Format konvertieren durchsucht und Datensätze herunterge-

📗 zim04f 📗 zim04g desy2007e … 📗 desy2007e\_res

Staging status 0.02 15-AUG-2012 15:06 Das Datenmanagementportal erlaubt auch Suchen über Ordnergrenzen hinweg, einfach anhand von Metadaten

## Analysieren

DESY betreibt eine Reihe von Hochleistungsrechnern. Darunter ist ein Rechencluster mit ca. 150 Knoten und Nvidia GPGPUs der Tesla-Serie.

Ein leistungsstarker HPC-Cluster ist mit 2500 Prozessorkernen, über 6 TB RAM, einem flinkem Netzwerk und großzügigem, extem schnellen lokalen Plattenplatz ausgestattet. Auf diesem Cluster können die Anwender massiv

parallele Berechnungen ausführen, inbesondere auch mit kommerziellen Anwendugen Mathematica oder Analysis

MATLAB.



Teilen



**Publication** 

Um auch sehr große Datenmengen einfach zwischen Instituten und Arbeitsgruppen teilen zu können, arbeiten die eu-Forschungszentren ropäischen NeXus, einem gemeinsamen Dateiformat.

dem einfacheren Dateiaustausch NeXus auch die Archivierung vereinfachen, da es, wie etwa auch JPG für Fotos oder MP3

einer Datei kombiniert.



Daten in klassischer Ordnerstruktur











Science für Musik, Metadaten und Nutzdaten in