

DAten aus PHotonen und Neutronen- Experimenten (DAPHNE-2)

Sprecher: Anton Barty, DESY, anton.barty@desy.de



Die Methoden der Synchrotron- und Neutronenstreuung werden in einer großen, interdisziplinären Bandbreite von Wissenschaftsfeldern angewendet. Die Nutzer repräsentieren dabei verschiedene Fachbereiche in den Naturwissenschaften, die sich dem gemeinsamen Bedarf an anspruchsvoller, schneller und tiefer Datenanalyse sowie den Herausforderungen der Implementierung eines qualifizierten Forschungsdatenmanagements gegenübersehen. Diese Nutzer von Photon- und Neutronengroßforschungseinrichtungen werden durch die gewählten Vertretungen "Komitee Forschung mit Synchrotronstrahlung" (KFS) und "Komitee Forschung mit Neutronen" (KFN) repräsentiert. KFS und KFN arbeiten bereits seit vielen Jahren zusammen, um die gemeinsamen Herausforderungen durch die digitale Transformation zu bewältigen. Ziel von DAPHNE ist es, eine übergreifende Infrastruktur zu schaffen, welche die Forschungsdaten entsprechend den FAIR-Prinzipien verarbeitet. Dazu bringt DAPHNE gezielt sowohl Nutzer als auch Betreiber der Großforschungseinrichtungen im Bereich Forschungsdatenmanagement zusammen.

Das DAPHNE-Konsortium wird das Forschungsdatenmanagement der mehr als 5000 deutschen Nutzer transformativ verändern und gleichzeitig, durch die etablierte europäische Vernetzung der Forschungsinfrastrukturen (RIs), ebenso auf die internationalen Anwender und Zentren ausstrahlen. Die Anwendungsgebiete innerhalb von DAPHNE sind dabei sehr vielfältig. In einer Umfrage des KFS gaben die Röntgennutzer ihr jeweiliges Fachgebiet mit Festkörperphysik (34 %), Oberflächen/Grenzflächen (18 %), biologische Systeme (14 %), weiche Materie (10 %), Atome/Moleküle/Gase (8 %), Flüssigkeiten (7 %), Optik / Laser (5%) oder der warmen dichten Materie (1 %) an. Ein ähnliches Bild ergibt sich für die Neutronennutzer (siehe Abb. 1). Die Bandbreite der untersuchten Themen und das Spektrum der eingesetzten Techniken sind daher sehr vielfältig; dementsprechend wird DAPHNE hauptsächlich die folgenden Disziplinen betreffen: Physik, Chemie, Biologie, Material- und Ingenieurwissenschaften, Umwelt- und Lebenswissenschaften.

An den von DAPHNE vertretenen Photon- und Neutronenzentren werden jährlich tausende von individuellen Nutzerexperimenten mit einer Vielzahl von Methoden durchgeführt. Einzelne Experimente können je nach Konfiguration Millionen von Dateien und teilweise über 700 TB Daten pro Woche erzeugen. Die erhobenen Daten sind dabei äußerst vielfältig, je nach Methode, Konfiguration der Beamline und Anwendungsbereich des Experimentes. Dabei ist die Menge der Daten sowie die Datenraten in den letzten Jahren rasant angestiegen, einerseits durch Verbesserung der Quellen, andererseits durch die Weiterentwicklungen moderner Detektoren.

Die Forschung im Bereich der Röntgen- und Neutronenstreuung steht daher sowohl in der Datenverarbeitung als auch in Speicherung und Management der Messdaten einer enormen Herausforderung gegenüber. Dies betrifft nicht nur die IT-Infrastruktur zur Datenspeicherung, sondern auch die Entwicklung neuer Algorithmen und Software-Konzepte zur Datenverarbeitung während und nach dem Experiment. Insbesondere das bisherige Vorgehen, bei dem die Wissenschaftler/innen ihre Rohdaten zur späteren Analyse mit in die Heimatinstitute nahmen, stößt zunehmend an seine Grenzen. Diese Veränderungen im Datenmanagement

und in der Datenanalyse erfordern eine sorgfältige Einbeziehung der damit verbundenen technischen, wissenschaftlichen und auch datenrechtlichen Implikationen. Die erfolgreiche Gestaltung der digitalen Transformation bedarf somit eines effizienten und nachhaltigen Forschungsdatenmanagements für die Nutzerschaft. Dieser Herausforderung stellt sich das DAPHNE Konsortium. Zudem birgt die gemeinsame Weiterentwicklung und Zusammenführung von existierenden Datenanalyse-Tools sowie der Austausch der Daten zwischen benachbarten NFDI Konsortien ein großes Potential in der Nachnutzung der Forschungsdaten.

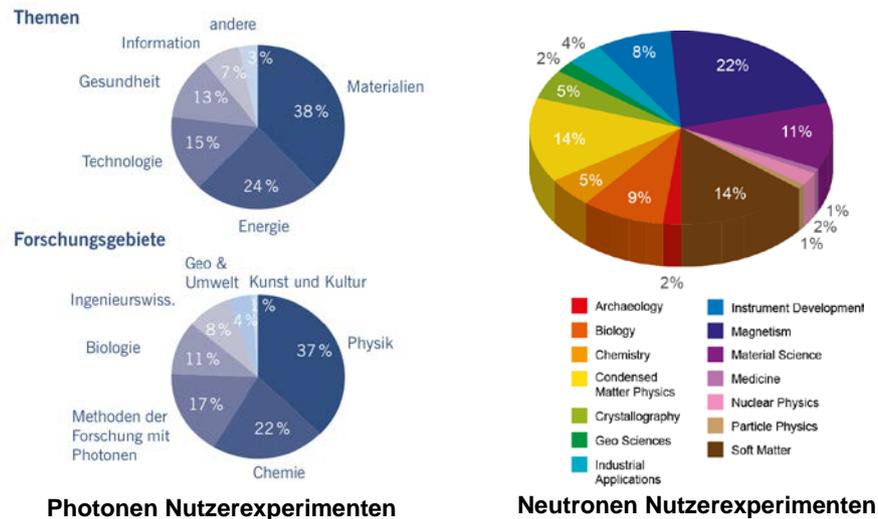


Abbildung 1: Typische sozio-ökonomische Herausforderungen und wissenschaftliche Bereiche, die von der KFS-Gemeinschaft angegangen werden (links: KFS-Umfrage 2018/19). Für Neutronen werden die Anwendungsgebiete der MLZ-Strahlzeittage bzw. der durchgeführten Experimente gezeigt (rechts: MLZ-Review 2013-2017).

DAPHNE bringt Großforschungseinrichtungen und Nutzer/innen aus den wichtigsten Anwendungsbereichen zusammen, um das Datenmanagement im Sinne der FAIR-Kriterien voranzutreiben. Dazu sollen folgende Maßnahmen durch DAPHNE realisiert und als Dienstleistung für die Photonen- und Neutronennutzergemeinschaft zur Verfügung gestellt werden:

1. Entwicklung effizienter Metadatendefinitionen, in Abstimmung mit anderen verwandten Wissenschaftsbereichen und anderen NFDI-Konsortien.
2. Information und Weiterentwicklung der Nutzergemeinschaft und Implementierung des FAIR- Datenmanagements und ihrer Anwendung.
3. Entwicklung eines elektronischen Laborbuches, das mit der Datenerfassung und Instrumentsteuerung verknüpft ist und den Bedürfnissen der Nutzer/innen an technisch effiziente Datenerhebung ebenso entspricht wie den Erfordernissen der Experimente. Im Idealfall erlaubt ein solches Laborbuch den reibungslosen Wechsel von einem Instrument oder Einrichtung zur anderen und stellt zudem die Grundlage für die Nachnutzung der Daten dar.
4. Entwicklung, Pflege und Bereitstellung von Nutzer-Analyse Software auf der Infrastruktur der Großforschungseinrichtungen zur gemeinsamen Softwareentwicklung, Etablierung von *Best-Practice* Verfahren und zur Sicherstellung der Reproduzierbarkeit von Analysemethoden.
5. Etablierung von Datenbanken und Katalogen für Anwenderexperimente, Proben, experimentelle Daten und Berechnungen, Referenzdatenbank und Analyse-Software

für veröffentlichte Ergebnisse, die durch PID / DOI mit der Rohdatenquelle verknüpft werden, um die Wiederverwendung von Daten und die gemeinsame Nutzung von Software innerhalb der Gemeinschaft und darüber hinaus zu ermöglichen.

DAPHNE knüpft dabei ganz wesentlich an die jahrelange Erfahrung in wissenschaftlicher Softwareentwicklung und Datenverarbeitung, die sowohl bei den Hochschulpartnern aber auch insbesondere bei den Großforschungseinrichtungen vorhanden ist. So spielt z.B. die IT Gruppe des DESY eine führende Rolle bei der Entwicklung des Softwarepakets dCache für das Datenmanagement (siehe dCache.org) einschließlich moderner Methoden der Qualitätssicherung, des Release-Managements und des Supports. MLZ ist der Hauptentwickler der Instrumentensteuerungssoftware NICOS (siehe nicos-controls.org). Die Erfahrungen und Verfahren aus diesen Projekten wie Code Peer Review, Continuous Integration Deployment Testing Cycles, Time-based Release und Bildung von Teams, die für die Programmierung und Release-Prozesse zuständig sind, werden von DAPHNE übernommen und genutzt werden. EMBL wird den Wissenstransfer zwischen Experten aus dem Bereich MX-Datenhandling und anderen Fachbereichen innerhalb DAPHNEs unterstützen. Zentren wie DESY, MLZ, EMBL und BESSY bringen große personelle Ressourcen mit in das DAPHNE Projekt ein, die besonders in der Form von Unterstützung bei der Software-Entwicklung und Bereitstellung eingesetzt werden. Dadurch wird bei DAPHNE eine hohe IT Qualität und professionelle Entwicklung bei gleichzeitiger Einbindung der Nutzerschaft sichergestellt.

DAPHNE ist eng mit europäischen Projekten wie z.B. der Photon und Neutron Open Science Cloud (PaNOSC) (finanziert durch den INFRAEOSC-04-Call) verbunden, die das Ziel haben FAIR Daten in sechs europäischen Forschungsinfrastrukturen zu realisieren. Weiterhin sind KFS und KFN jeweils an die Europäischen Nutzervertretungen angebunden (ESUO - European Synchrotron User Organization, ENSA - European Neutron Scattering Association) und kooperieren mit den Vertretungen der Facilities - League of Advanced European Neutron Sources, LENS, und LEAPS - League of Advanced Photon Sources.

Das Konsortium DAPHNE repräsentiert die breite Nutzergemeinschaft, die Röntgen- und Neutronentechniken einsetzt, und die bundesweit über 5000 Wissenschaftler/innen umfasst. Traditionell interagieren Großforschungseinrichtungen und Nutzer/innen sehr eng bei der Entwicklung von Methoden und Datenaufnahmetechniken. Da die Implementierung eines effizienten Forschungsdatenmanagements unbedingt einen gemeinsamen und kohärenten Ansatz erfordert, ist diese umfassende Zusammenarbeit der zentrale Bestandteil von DAPHNE. Das Konsortium umfasst die Zentren und Nutzer/innen an Großforschungsanlagen wie die Röntgenquellen PETRA III und FLASH bei DESY, BESSY II beim HZB, den European-XFEL, das ELBE-ZENTRUM für Hochenergiestrahlungsquellen im HZDR und der ESRF in Grenoble, sowie die Neutronenquellen im Heinz Maier Leibnitz Zentrum und im Institut Laue-Langevin in Grenoble, sowie zukünftige Quellen wie PETRA IV, BESSY III und die European Spallation Source.

Durch die breite Zusammensetzung der Photonen- und Neutronennutzerschaft besitzt DAPHNE Anknüpfungspunkte in viele Fachgebieten - insbesondere im Bereich der Metadatendefinitionen und der Workflows. Somit können die in DAPHNE erarbeiteten Konzepte für Datenmanagement und die Etablierung von FAIR Daten auch als Vorbild für die NFDI über Fachdisziplinen hinweg dienen.

Vorgesehene Mitglieder des Konsortiums (Co-Sprecherinnen/Co-Sprecher und die weiteren, beteiligten Institutionen):

Co-Sprecher/in	Zugehörige Institution
Dr. Anton Barty Head of photon science computing group anton.barty@desy.de	Deutsches Elektronen-Synchrotron DESY Photon Science Notkestrasse 85 22607 Hamburg
Prof. Dr. Christian Gutt Professor for Experimental and X-ray Physics gutt@physik.uni-siegen.de	Universität Siegen Department Physik Walter-Flex-Strasse 3 ENC B016 57072 Siegen
Dr. Astrid Schneidewind Head of PANDA TAS group/Senior Scientist Chair of KFN a.schneidewind@fz-juelich.de	Jülich Centre for Neutron Science JCNS Forschungszentrum Jülich GmbH JCNS at MLZ Garching Lichtenbergstraße 1 85747 Garching
Dr. Wiebke Lohstroh Gruppenleitung Instrumente Wiebke.lohstroh@frm2.tum.de	Forschungsneutronenquelle Heinz Maier-Leibnitz (FRM II) Technische Universität München Lichtenbergstr. 1 85748 Garching
Prof. Dr. Jan-Dierk Grunwaldt Technische Chemie und Polymerchemie grunwaldt@kit.edu	Karlsruher Institut für Technologie Institut für Technische Chemie und Polymerchemie Engesserstr. 20 76131 Karlsruhe
PD Dr. Bridget Murphy LISA group leader Chair of KFS murphy@physik.uni-kiel.de	Ruprecht Haensel Labor Institut für Experimentelle und Angewandte Physik Christian-Albrechts-Universität zu Kiel Olshausenstr. 40 24098 Kiel
Dr. Sebastian Busch Deputy head of HZG MLZ outstation sebastian.busch@hzg.de	Helmholtz-Zentrum Geesthacht Zentrum für Material- und Küstenforschung GmbH Max-Planck-Straße 1 21502 Geesthacht
Prof. Dr. Tobias Unruh Professur für Nanomaterialcharakterisierung Tobias.Unruh@fau.de	Nanomaterialcharakterisierung (Streuethoden) Friedrich-Alexander-Universität Erlangen-Nürnberg Lehrstuhl für Kristallographie und Strukturphysik Staudtstraße 3 91058 Erlangen
Prof. Dr. Frank Schreiber Professor of Experimental Physics frank.schreiber@uni-tuebingen.de	Institut für Angewandte Physik Universitaet Tuebingen Auf der Morgenstelle 10 72076 Tuebingen
Dr. Thomas R. Schneider Joint Head of Research Infrastructures thomas.schneider@embl-hamburg.de	European Molecular Biology Laboratory c/o DESY Notkestr. 85 22607 Hamburg
Dr. Jörg Hammel Beamline Scientist/Senior Scientist joerg.hammel@hzg.de	Helmholtz-Zentrum Geesthacht Zentrum für Material- und Küstenforschung GmbH Max-Planck-Straße 1 21502 Geesthacht

Prof. Dr. Hans Fangohr Head of Data Analysis hans.fangohr@xfel.eu	European XFEL GmbH Holzkoppel 4 22869 Schenefeld
Dr. Frank Weber Group leader Frank.Weber@kit.edu	Karlsruher Institut für Technologie Institut für QuantenMaterialien und Technologien Hermann-von-Helmholtz-Platz 1 76344 Eggenstein-Leopoldshafen
Prof. Dr. Birgit Kanngießner Stiftungsprofessur Analytische Röntgenphysik birgit.kanngiesser@tu-berlin.de	Analytische Röntgenphysik Institut für Optik und Atomare Physik Technische Universität Berlin 10623 Berlin
Dirk Lützenkirchen Hecht Group leader dirklh@uni-wuppertal.de	University of Wuppertal School of Mathematics and Natural Sciences Gauß-Str. 20 42097 Wuppertal
Prof. Dr. Sarah Köster Professor of Experimental Physics sarah.koester@phys.uni-goettingen.de	Georg-August-Universität Göttingen Institute for X-Ray Physics Friedrich-Hund-Platz 1 37077 Göttingen
Prof. Dr. Paola Coan Professor for Experimental Medical Physics paola.coan@physik.uni-muenchen.de	Ludwig Maximilians University Department of Physics Am Coulombwall 1 85748 Garching /Faculty of Medicine Dept. Clinical Radiology Marchioninstr. 15 81377 Großhadern - München
Dr. Andreas Houben (Dipl.-Chem.) Senior Scientist andreas.houben@ac.rwth-aachen.de	RWTH Aachen Institut für Anorganische Chemie Landoltweg 1A 52074 Aachen
Heike Görzig Scientific /computing staff scientist heike.goerzig@helmholtz-berlin.de	Research Data Management Group Helmholtz-Zentrum Berlin für Materialien und Energie Hahn-Meitner-Platz 1 14109 Berlin
Dr. Ingo Manke Head of Imaging Group manke@helmholtz-berlin.de	Helmholtz-Zentrum Berlin für Materialien und Energie Hahn-Meitner-Platz 1 14109 Berlin