

HELMHOLTZ INFRASTRUCTURE FOR FEDERATED ICT SERVICES (HIFIS)



INHALTSVERZEICHNIS

| | |
|--|-----------|
| EINLEITUNG | 3 |
| 1 ZUSAMMENFASSUNG IN DEUTSCH UND ENGLISCH | 9 |
| 1.1 Zusammenfassung | 9 |
| 1.2 Summary | 10 |
| 2 EINLEITUNG, MOTIVATION | 11 |
| 3 ZIELE | 18 |
| 4 STRUKTUR DER AKTIVITÄTEN | 20 |
| 5 ARBEITSPROGRAMM | 22 |
| 5.1 Cloud Services | 22 |
| 5.2 Backbone Services | 33 |
| 5.3 Software Services | 39 |
| 6 ARBEITSPAKETE UND MEILENSTEINE | 46 |
| 7 GOVERNANCE | 47 |
| 8 PERSONELLE UND FINANZIELLE RESSOURCEN | 50 |
| 9 ANHANG | 52 |
| 9.1 Abkürzungsverzeichnis | 52 |
| 9.2 Übersicht über die Planung der Aufbauphase | 55 |
| 9.3 HIFIS Verortungsvorschlag | 56 |
| 9.4 Umfeldanalyse | 58 |
| 9.5 SWOT-Analyse | 64 |
| 9.6 Zusammenarbeit mit anderen Plattformen und externen Partnern | 67 |
| 9.7 Bibliografie | 67 |
| DANKSAGUNG | 71 |

EINLEITUNG IN INFORMATION & DATA SCIENCE IN DER HELMHOLTZ-GEMEINSCHAFT

Information & Data Science – Eine globale Herausforderung

Die digitale Transformation ist eine der größten Herausforderungen für Wissenschaft, Wirtschaft und Gesellschaft zu Beginn des 21. Jahrhunderts. Sie bietet enorme Chancen in nahezu allen Bereichen des Lebens – innovative Formen von Arbeit und Zusammenleben, völlig neuartige Plattformen für Handel und Wissenschaft, ungeahnte Möglichkeiten für die Medizin, revolutionäre Ansätze für Netzwerke der Energieversorgung, zukunftsweisende Verfahren für Klimaschutz und vieles mehr.

Im Kern der digitalen Transformation steht die Wertschöpfungs- und Erkenntniskette „von Daten zu Wissen zu Innovation“. Dies ist auch die entscheidende Herausforderung in allen wissenschaftlichen Disziplinen, welche die Grundlagen für den gesellschaftlichen Fortschritt der Zukunft bereiten.

Damit Deutschland diese Entwicklung mitgestalten und daran partizipieren kann, sind ambitionierte, mutige und zukunftsweisende Schritte erforderlich. Die Verbindung von Informatik, Mathematik, Statistik, Sensortechnologie, Simulation und datenintensivem Rechnen mit anspruchsvollen Anwendungsfeldern aus dem breiten Spektrum der Natur- und Ingenieurwissenschaften, der Medizin sowie den Geistes- und Sozialwissenschaften wird ein dynamischer Motor des Innovations- und Forschungsstandortes Deutschland sein. Diese innovativen Verbindungen werden völlig neue wissenschaftliche Erkenntnisse mit erheblichen Mehrwerten für Wirtschaft und Gesellschaft hervorbringen.

Deutschland hat im Feld der Hochtechnologien mit großem transformativem Potenzial eine sehr gute Ausgangsposition im weltweiten Vergleich. Um diese Position zu halten und auszubauen, sind neben gesamtgesellschaftlichen Anstrengungen auch völlig neuartige Ansätze im Wissenschaftssystem erforderlich.

Information & Data Science in der Helmholtz-Gemeinschaft

Die Helmholtz-Gemeinschaft Deutscher Forschungszentren leistet Beiträge zur Lösung großer und drängender Fragen von Gesellschaft, Wissenschaft und Wirtschaft durch interdisziplinäre, wissenschaftliche Spitzenleistungen in sechs Forschungsbereichen: Energie, Erde und Umwelt, Gesundheit, Schlüsseltechnologien, Materie sowie Luftfahrt, Raumfahrt und Verkehr. Sie ist mit über 39.000 Mitarbeiterinnen und Mitarbeiter in 18 Forschungszentren die größte Wissenschaftsorganisation Deutschlands.

Als Forschungsorganisation, die sich zum Ziel gesetzt hat große und aktuelle gesellschaftliche Herausforderungen zu adressieren, hat die Helmholtz-Gemeinschaft in den letzten Jahrzehnten ein enormes Kompetenzportfolio im Bereich Information & Data Science aufgebaut: auf Gebieten wie Informationsverarbeitung, Big Data, Datenanalyse, Simulation, Modellierung, Bioinformatik, bildgebenden Verfahren, Forschungsdaten-Management, High Performance Computing, Robotik, technischen sowie biologischen Informationssystemen und vielen weiteren zukunftsweisenden Technologien. Wie kaum eine andere Forschungsorganisation verfügt sie über einen exponentiell wachsenden Schatz von Big Data.

Die Helmholtz-Zentren und Forschungsbereiche haben auch eine herausragende Ausgangsposition für eine erfolgreiche, synergetische Verbindung der Kompetenzen: als Betreiber großer Forschungsinfrastrukturen (beispielsweise Satellitenmissionen oder Großanlagen der Kern- und Teilchenphysik), als Anwender von Supercomputing der neuesten Generation sowie als Kompetenzträger für komplexe Simulationen (beispielsweise umfangreicher Erd- und Klimamodelle, virtuelles Materialdesign und Systembiologie). Sie sind daher auch

schon seit Langem mit dem gesamten Data-Lifecycle vertraut: der Forschungsplanung, dem Erheben, der Handhabung und Pflege, dem Analysieren, dem Auswerten und der Nutzbarmachung sehr großer und komplexer Datenmengen. An allen Standorten und in allen Einzeldisziplinen gibt es teilweise weltweit führende Ansätze und herausragende Methodenkompetenz – aber gerade auch in ihrem Zusammenspiel ergeben sich ungeahnte Möglichkeiten.

Darauf aufbauend und darüber hinausgehend verstärkt die Helmholtz-Gemeinschaft im hochaktuellen Themenfeld Information & Data Science die eigene Kompetenz, schafft Synergien in der Forschungslandschaft, greift die Entwicklungen im nationalen, europäischen und internationalen Kontext auf und setzt neue disruptive Ansätze um. Neben der Stärkung der Einzeldisziplinen mit modernsten daten- und informationswissenschaftlichen Methoden ist es erklärtes Ziel, das Thema Information & Data Science auf Gemeinschaftsebene disziplinübergreifend voranzutreiben.

Die Helmholtz-Gemeinschaft adressiert das komplexe Themenfeld Information & Data Science auf allen Ebenen:

- Es hat hohe Priorität in der Agenda des Präsidenten.
- Helmholtz-Zentren und Forschungsbereiche berücksichtigen in ihren jeweiligen Strategien das große Potenzial dieses Themenfeldes.
- Im Rahmen der Neuausrichtung der Forschungsbereiche wird der bisherige Forschungsbereich Schlüsseltechnologien in einen neuen Forschungsbereich Information weiterentwickelt.
- Die Gemeinschaft errichtete mehrere neue Institute zu Simulations- und Datentechnologien und Cyber-sicherheit und plant aktuell die Aufnahme eines neuen Helmholtz-Zentrums für Informationssicherheit am Standort Saarbrücken.
- Wissenschaftlicher Nachwuchs wird in diesem Bereich auf neuen Wegen und in großer Zahl in einem neuen Netzwerk regionaler Helmholtz Information & Data Science Schools (HIDSS) ausgebildet.
- Der von der Helmholtz-Gemeinschaft initiierte Helmholtz-Inkubator Information & Data Science potenziert die einzelnen Initiativen als ein neuartiger, gemeinschaftsweiter Think-Tank und Zukunftsmotor.
- Durch Ihre intensive Interaktion mit nationalen und internationalen Partnern leistet Helmholtz auch entscheidende Beiträge zu Initiativen wie der Nationalen Forschungsdaten Infrastruktur (NFDI) und zu internationalen Allianzen im Forschungsdatenmanagement

Der Helmholtz-Inkubator Information & Data Science

Der Helmholtz-Inkubator Information & Data Science wurde im Juni 2016 vom Präsidenten der Helmholtz-Gemeinschaft ins Leben gerufen, um die vielfältige, dezentrale Expertise der Gemeinschaft im weiten Themenfeld Information & Data Science intelligent zusammenzuführen und zu potenzieren.

Dazu haben alle Helmholtz-Zentren je zwei hochkarätige Wissenschaftlerinnen und Wissenschaftler in den Helmholtz-Inkubator entsandt. Diese 36 Fachleute vertreten zusammen eine enorme fachliche Breite und decken viele innovative Kompetenzen ab. Unterstützt werden sie durch ausgewiesene Expertinnen und Experten aus forschenden Unternehmen und namhaften Forschungseinrichtungen sowie mehreren Beratungsunternehmen. Begleitet wird der Helmholtz-Inkubator von der Geschäftsstelle der Helmholtz-Gemeinschaft.

Der Helmholtz-Inkubator integriert bestehende, zukunftsweisende Initiativen der Helmholtz-Gemeinschaft durch einen gemeinschaftsweiten Bottom-up-Prozess. Die regelmäßige Zusammenführung und Verdichtung

der Expertise der Helmholtz-Zentren ermöglicht die visionäre Gestaltung des Themas Information & Data Science über die Grenzen von Zentren und Forschungsbereichen hinaus. Dabei geht der Helmholtz-Inkubator völlig neue Wege, um das Zukunftsfeld dynamisch, umfangreich und fachübergreifend durch Setzung strategischer Schwerpunkte zu gestalten.

Der Helmholtz-Inkubator verfolgt derzeit folgende Ziele:

- die regelmäßige Interaktion kreativer Köpfe aus der gesamten Gemeinschaft,
- die Schaffung von Grundlagen für innovative, interdisziplinäre Netzwerke und Ansätze,
- die Identifizierung zukunftsweisender Themenfelder und disruptiver Pilotprojekte,
- die Planung und Begleitung von langfristig angelegten und gemeinschaftsweiten Plattformen.

Der Helmholtz-Inkubator hat innerhalb der Helmholtz-Gemeinschaft eine hohe Dynamik entfaltet, zahlreiche Impulse mit großer thematischer Breite formuliert, vielfältigen Austausch zwischen Digitalisierungsexpertinnen und -experten aus allen Domänen ermöglicht und so einen tiefgehenden Strategieprozess eingeleitet, der in dieser Form und in diesem Umfang einmalig im deutschen Wissenschaftssystem ist.

Über 150 beteiligte Wissenschaftlerinnen und Wissenschaftler, davon 36 Vertreterinnen und Vertreter des Helmholtz-Inkubators, sowie 10 externe Beraterinnen und Berater (u.a. IBM, SAP, Trumpf, Blue Yonder, Gauß-Allianz) haben sich in den letzten zwei Jahren im Rahmen von über 35 Workshops und AG-Treffen diesem großen Themenkomplex gewidmet. Auch die Vorstände der Helmholtz-Gemeinschaft haben sich in einem dedizierten Workshop mit den strategischen Überlegungen des Helmholtz-Inkubators auseinandergesetzt. Nach Schätzungen der Geschäftsstelle sind bis heute über 25.000 Personenstunden Arbeit in diesen Prozess geflossen.

Der bisherige Prozess

Der Helmholtz-Inkubator trat im Oktober 2016 zu seinem ersten, zweitägigen Workshop zusammen und diskutierte das Thema Information & Data Science offen und uneingeschränkt; jedes Zentrum lieferte Impulse aus seiner jeweiligen Sicht.

Obwohl zu Beginn des Prozesses die Schaffung von Konsortien zur Bearbeitung von Pilotprojekten im Vordergrund stand, formulierten die Expertinnen und Experten die klare Empfehlung, dass die Entwicklung von dauerhaften, strukturellen Aktivitäten im Themenbereich Information & Data Science auf einigen Teilgebieten auf Gemeinschaftsebene notwendig sei – zusätzlich zu neuen Impulsen durch Pilotprojekte.

Die Vertreterinnen und Vertreter erörterten die zentralen Herausforderungen der Helmholtz-Gemeinschaft auf Ebene der gesamten Gemeinschaft. Die identifizierten Themen wurden umfassend mit allgemeinen forschungspolitischen Initiativen der Allianz-Organisationen und Vorstellungen der Zuwendungsgeber sowie internationalen Entwicklungen abgeglichen. Zusätzlich wurde der Helmholtz-Inkubator dabei von führenden externen Expertinnen und Experten im Thema Information & Data Science beraten.

Die Vorschläge zu gemeinschaftsweit zu bearbeitenden Themenfeldern wurden im Frühjahr 2017 in mehreren fokussierten Arbeitsgruppen des Helmholtz-Inkubators konkretisiert. Die Arbeitsgruppen setzen sich aus den Inkubator-Mitgliedern zusammen, die sich spezifisch in diesem Thema einbringen wollen; weitere mit den Themen befasste Expertinnen und Experten der Helmholtz-Gemeinschaft nahmen beratend an den Sitzungen der Arbeitsgruppen teil. Die Ergebnisse der Arbeitsgruppen wurden in einem Inkubator-Workshop mit allen Vertreterinnen und Vertretern des Helmholtz-Inkubators und externen Expertinnen und Experten im Mai 2017 diskutiert und weiter geschärft.

Aus den diskutierten Themenvorschlägen wurden diejenigen Themen weiter verfolgt, von denen gemeinschaftsweite Mehrwerte und ein erhebliches Entwicklungspotenzial erwartet werden konnten. Der Helmholtz-Inkubator identifizierte so fünf Themenkomplexe, die er der Helmholtz-Gemeinschaft zur langfristigen Verfolgung durch die Etablierung von fünf dedizierten Plattformen vorschlug.

Die Vorstände der Helmholtz-Gemeinschaft haben sich in einer außerordentlichen, fachlichen Sitzung im September 2017 eingehend mit diesen fünf Plattform-Ansätzen befasst. Zu allen Themen wurden inhaltliche und strukturelle Leitplanken für eine detaillierte Konzepterstellung formuliert. Auf der Mitgliederversammlung haben sie diese zur detaillierten Ausarbeitung durch den Helmholtz-Inkubator und seiner Arbeitsgruppen empfohlen, um eine potenzielle Umsetzung ab September 2018 vorzubereiten. Die Mitgliederversammlung forderte die Geschäftsstelle dazu auf, die AGs bei der Ausgestaltung von Governance und Finanzierungsfragen zu unterstützen.

Die Ergebnisse des Vorstandsworkshop Information & Data Science und der Mitgliederversammlung wurden anschließend dem Senat der Helmholtz-Gemeinschaft übermittelt.

Die Arbeitsgruppen des Helmholtz-Inkubators haben darauf folgend ab Oktober 2017 die weitere Konkretisierung der Themen vorgenommen. Die Inkubator-Vertreter der Zentren wurden nochmals aufgefordert, spezialisierte Fachleute in die jeweiligen AGs zu entsenden; so wurde themenspezifisch ein noch höheres Maß an Expertise aufgebaut.

Die Arbeitsgruppen haben zur weiteren Ausarbeitung der Konzepte Berichterstatter bestimmt. Zur Unterstützung der Arbeitsgruppen und Berichterstatter wurden pro Themenfeld zwei Referenten eingestellt; hierfür stellte der Präsident Mittel aus dem Impuls- und Vernetzungsfonds bereit. Die AGs bildeten folgende Teams:

- Arbeitsgruppe wissenschaftlicher Nachwuchs: vertreten durch die Berichterstatter Achim Streit (KIT) und Uwe Konrad (HZDR), 28 Mitarbeitende, unterstützt durch die Projektreferentin Susan Trinitz. Erarbeitetes Konzept: Helmholtz Information & Data Science Academy (HIDA).
- Arbeitsgruppe Mehrwerte aus Forschungsdaten durch Metadaten: vertreten durch die Berichterstatter Rainer Stotzka (KIT), Kirsten Elger (GFZ) und Frank Ückert (DKFZ), 27 Mitarbeitende, unterstützt durch die Projektreferentinnen Nanette Reißler-Pipka und Romyana Proynova. Erarbeitetes Konzept: Helmholtz Metadata Center (HMC).
- Arbeitsgruppe Basistechnologien und grundlegende Dienste: vertreten durch die Berichterstatter Ants Finke (HZB), Volker Gülzow (DESY) und Uwe Konrad (HZDR), 36 Mitarbeitende, unterstützt durch die Projektreferenten Knut Sander und Tobias Frust. Erarbeitetes Konzept: Helmholtz Infrastructure for Federated ICT Services (HIFIS).
- Arbeitsgruppe bildgebende Verfahren (Imaging): vertreten durch die Berichterstatter Christian Schroer (DESY) und Wolfgang zu Castell (HMGU), 44 Mitarbeitende, unterstützt durch die Projektreferenten Alexander Pichler und Murali Sukumaran. Erarbeitetes Konzept: Helmholtz Imaging Platform (HIP).
- Arbeitsgruppe Künstliche Intelligenz und Maschinelles Lernen: vertreten durch die Berichterstatter Fabian Theis (HMGU) und Katrin Amunts (FZJ), 45 Mitarbeitende, unterstützt durch die Projektreferentinnen Susanne Wenzel und Angela Jurik-Zeiller. Erarbeitetes Konzept: Helmholtz Artificial Intelligence Cooperation Unit (HAICU).

Im März 2018 fand der vierte Inkubator-Workshop statt. Hier wurden die bis zu diesem Zeitpunkt erarbeiteten Konzeptentwürfe präsentiert und weiterentwickelt. Dabei wurden die Inkubator-Vertreter erneut von ausge-

wiesenen externen Expertinnen und Experten unterstützt. Die Ergebnisse des Inkubator-Workshops wurden im Anschluss in den jeweiligen Arbeitsgruppen umgesetzt und die Konzepte von den Arbeitsgruppen finalisiert. Flankierend fanden individuelle Gespräche aller Berichterstatter mit der Geschäftsführerin der Helmholtz-Geschäftsstelle zu Fragen der Governance und Finanzierung statt.

Nach Fertigstellung der Konzepte durch die Arbeitsgruppen und Berichterstatter wurden diese im Juli 2018 allen offiziellen Inkubator-Vertretern übergeben. Der Helmholtz-Inkubator empfiehlt den Vorständen der Helmholtz-Gemeinschaft die vorliegenden fünf Konzepte zur nachhaltigen und gemeinschaftsweiten Umsetzung.

Die Vorstände der Helmholtz-Gemeinschaft können nun im Rahmen der Mitgliederversammlung im September 2018 über die Umsetzung und Finanzierung jedes Plattform-Konzepts befinden. Die Vorstände bestimmen außerdem für jedes genehmigte Konzept, welche Helmholtz-Zentren die Plattformen zum Nutzen der gesamten Helmholtz-Gemeinschaft tragen sollen oder definieren ein Verfahren zur Verortung der Plattformen.

Die wichtigsten Prozessschritte der vergangenen zwei Jahre sind in Abbildung E.1 schematisch dargestellt.

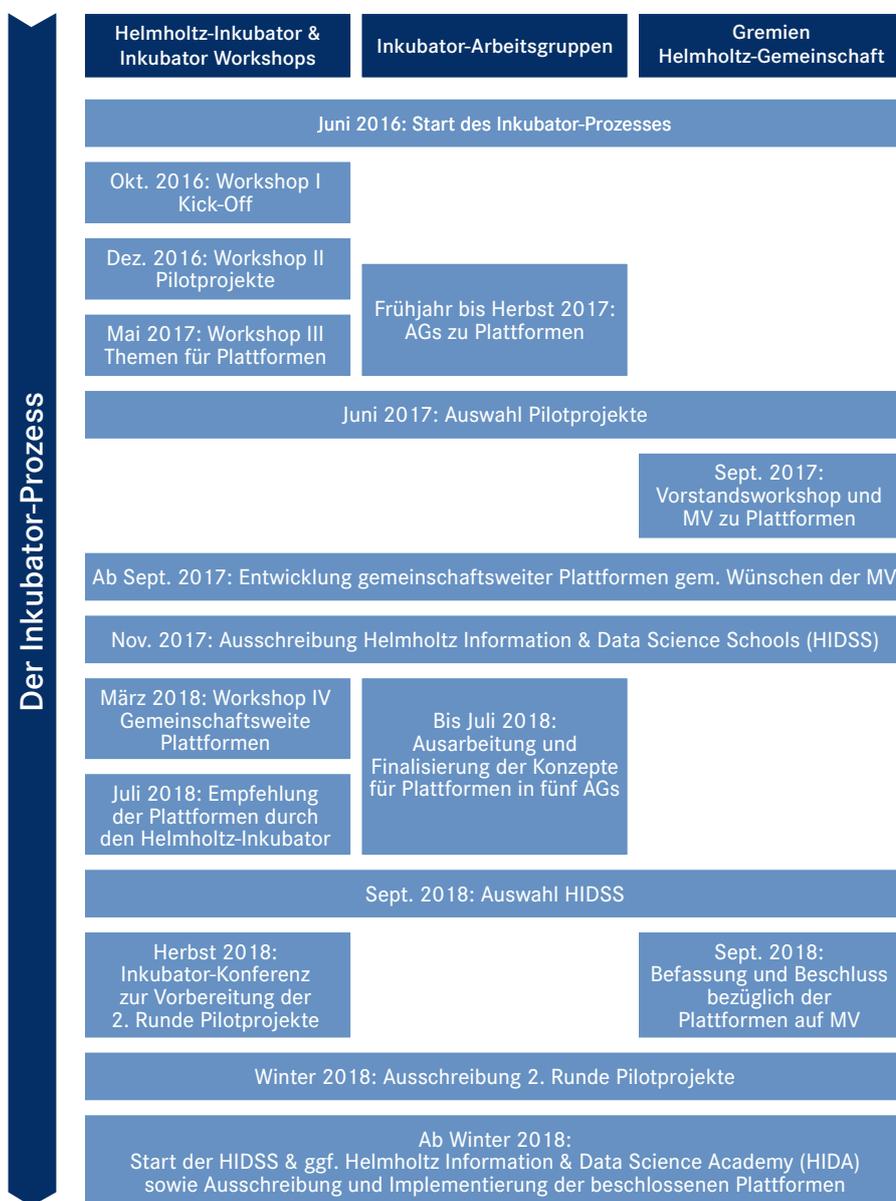


Abbildung E.1: Die wichtigsten Schritte des Inkubator-Prozesses der vergangenen zwei Jahre.

1 ZUSAMMENFASSUNG IN DEUTSCH UND ENGLISCH

1.1 ZUSAMMENFASSUNG

Die Helmholtz-Gemeinschaft mit mehr als 39.000 Mitarbeitern stellt ihren Wissenschaftlerinnen und Wissenschaftlern und auch den wissenschaftlichen Nutzern außerhalb der Helmholtz-Gemeinschaft ein hervorragendes Angebot an wissenschaftlichen Großgeräten und Forschungsinfrastrukturen zur Verfügung. Die Nutzung dieser Einrichtungen basiert dabei zu großen Teilen auch auf Informations- und Kommunikationstechnologien, die immer stärker die Arbeits- und Forschungsprozesse beeinflussen. Die Digitalisierung verändert bereits heute die Arbeitsweisen und Methoden in den Wissenschaften und wird zukünftig die Weiterentwicklung und Neugestaltung der Forschungsarbeit prägen. Diesen Anforderungen aus der Wissenschaft müssen die IT-Infrastrukturen der Zukunft gewachsen sein, damit die Helmholtz-Gemeinschaft Lösungen für die drängenden gesellschaftlichen Herausforderungen der Zukunft entwickeln kann. Dazu benötigen Wissenschaftler:

- ▶ eine nahtlose, Zentren-übergreifende IT-Infrastruktur mit integrierten ICT-Dienstleistungen auf der Basis schneller Netze und einheitlichem Nutzerzugang,
- ▶ einen in die Zusammenarbeits- und Forschungsprozesse integrierten sicheren, effizienten und weltweit verfügbaren Daten- und Anwendungszugriff auf der Basis von Cloud Diensten,
- ▶ Ausbildung und Unterstützung, um qualitativ hochwertige und nachhaltige Software zu entwickeln und zu veröffentlichen.

Die Forschung in den Helmholtz-Zentren entwickelt sich zunehmend Zentren-übergreifend, so bilden sich thematisch orientiert Cluster von Zentren und nutzen gemeinsam Forschungsinfrastrukturen wie Beamlines, Schiffe etc. Drei exemplarische Fälle werden vorgestellt, die von der zuverlässigen Speicherung von Beobachtungsdaten in Erde & Umwelt über standardisierte Analysepipelines und den Zugriff auf empirische Daten aus Kohortenstudien im Bereich Life Science bis hin zur Nutzung großer Instrumente wie der Helmholtz International Beamline am XFEL im Bereich Materie reichen.

Der vorliegende Vorschlag „Helmholtz Infrastructure for Federated ICT Services“ (HIFIS) zielt auf eine hervorragende Infrastruktur in der Informationstechnologie der Helmholtz-Gemeinschaft ab, um die Wissenschaft auf allen Ebenen zu fördern. HIFIS orientiert sich nicht nur auf Anforderungen größter Datenmengen oder Rechenleistungen. Vielmehr unterstützt HIFIS auch das breite Spektrum aller anderen Forschungsarbeiten. Es hilft, Helmholtz wettbewerbsfähig zu halten, im Vergleich zu Initiativen anderer Länder wie den USA oder Großbritannien, die den Ausbau ihrer wissenschaftlichen Einrichtungen auf dem Gebiet der ICT Services strategisch voranbringen. Die Schaffung einer herausragenden Informations-Infrastruktur ist die Voraussetzung, um als starke Gemeinschaft Partnerschaften mit Universitäten und Forschungseinrichtungen einzugehen. HIFIS erstellt Strukturen und Services, durch die sich die Helmholtz-Gemeinschaft in nationalen und internationalen Initiativen wie der Nationalen Forschungsdaten-Infrastruktur (NFDI) und der European Open Science Cloud (EOSC) erfolgreich positionieren kann.

Die Schlüsselemente dieses Vorschlags sind der Aufbau eines schnellen „Virtual Private Network“ (VPN) zwischen den Hauptstandorten der Zentren und die Installation von Basisdiensten wie Vertrauensbeziehungen und einer gemeinsamen Authentifizierungs- und Autorisierungs-Infrastruktur (AAI). Dieses Helmholtz-Backbone bildet die Basis für übergeordnete Dienste wie eine Helmholtz-Cloud, die in einigen Zentren schon vorhandene,

exzellente Services aufnimmt. Diese Helmholtz-Cloud bietet Datenaustausch- und -analysefunktionen sowie einen Ort der Kommunikation und der Zusammenarbeit. Hinzu kommt der Aufbau einer Plattform, die moderne Software Engineering Methoden und -Infrastrukturen bereitstellt und über Software-Repositoryen in Helmholtz-Zentren entwickelte Softwarelösungen auf einem definierten Qualitätsniveau sichtbar und nachnutzbar macht.

Dieser Vorschlag orientiert sich in erster Linie am Aufbau von IT-Infrastrukturen und den für die Forschung notwendigen Diensten. Dazu sind insbesondere personelle Ressourcen notwendig, um diese Dienste zu entwickeln, einzurichten und zu betreiben. Er knüpft u. a. an Komponenten der Hardware Ressourcen an, die derzeit in der Helmholtz-Data Federation (HDF) installiert werden und schließt deren Arbeit an den Diensten ein. Ziel dieses Vorhabens ist es, in die Breite der Helmholtz-Zentren zu gehen. Deshalb wurde von vornherein ein förderierter Ansatz gewählt, der die Zentren einbindet, die bereits über erhebliche Kompetenzen auf den jeweiligen Gebieten verfügen.

1.2 SUMMARY

With more than 39,000 employees, the Helmholtz Association provides its scientists and scientific users outside the Helmholtz Association with an excellent range of large scientific facilities and research infrastructures. The use of these facilities is largely based on information and communication technologies, which are increasingly influencing work and research processes. Digitisation is already changing the working methods and methods in the sciences and will shape the further development and redesign of research work in the future. The IT infrastructures of the future must meet these scientific requirements so that the Helmholtz Association can develop solutions for the urgent social challenges of the future. For this scientists require

- ▶ a seamless, cross-centre IT infrastructure with integrated ICT services based on fast networks and uniform user access,
- ▶ a secure, efficient and globally available data and application access based on cloud services integrated into collaboration and research processes,
- ▶ training and support to develop and publish high quality and sustainable software.

Research at the Helmholtz Centres is increasingly developing across centres, forming thematically oriented clusters of centres and using research infrastructures such as beamlines, ships etc. together. Three exemplary cases are presented, ranging from the reliable storage of observation data in Earth & Environment, standardised analysis pipelines and access to empirical data from cohort studies in the field of life science to the use of large instruments such as the Helmholtz International Beamline at the XFEL in the field of matter.

The present proposal „Helmholtz Infrastructure for Federated ICT Services“ (HIFIS) aims at an excellent infrastructure in the information technology of the Helmholtz Association in order to promote science at all levels. HIFIS is not only oriented to the requirements of the largest amounts of data or computing power. HIFIS also supports the broad spectrum of all other research work. It helps to keep Helmholtz competitive in comparison to initiatives of other countries such as the USA or Great Britain, which strategically advance the expansion of their scientific institutions in the field. The creation of an outstanding information infrastructure is a prerequisite for entering into partnerships with universities and research institutions as a strong community. HIFIS creates structures and services through which the Helmholtz Association can successfully position itself in national and international initiatives such as the National Research Data Infrastructure (NFDI) and the European Open Science Cloud (EOSC).

The key elements of this proposal are the establishment of a fast Virtual Private Network (VPN) between the main sites of the centres and the installation of basic services such as trust relationships and a common authentication and authorisation infrastructure (AAI). This Helmholtz backbone forms the basis for higher-level services such as a Helmholtz Cloud, which in some centres incorporates already existing, excellent services. This Helmholtz Cloud offers data exchange and analysis functions as well as a place for communication and cooperation. In addition, a platform is being set up to provide modern software engineering methods and infrastructures and to make software solutions developed via software repositories in Helmholtz Centres visible and usable at a defined quality level.

This proposal is primarily oriented towards the development of IT infrastructures and the services required for research. This requires personnel resources in particular to develop, set up and operate these services. It ties amongst other things to the hardware resources currently installed throughout the Helmholtz Data Federation (HDF) and includes their work on the services. The aim of this project is to expand into the Helmholtz Centres. For this reason, a federated approach was chosen from the outset, involving the centres that already have considerable expertise in their respective fields.

2 EINLEITUNG, MOTIVATION

Die Helmholtz-Gemeinschaft steht im Wettbewerb um die besten Köpfe. Eine herausragende ICT Infrastruktur ist für die Gewinnung ehrgeiziger und hochqualifizierter Forscherinnen und Forscher für die Helmholtz-Gemeinschaft ein sehr bedeutsames Kriterium. Wissenschaftsorganisationen in anderen Ländern wie z. B. die National Science Foundation (NSF) in den USA, das European Bioinformatics Institute (EMBL EBI), das Science and Technology Facilities Council (STFC) in UK oder das Centre national de la recherche scientifique (CNRS) in Frankreich haben eine hoch performante Vernetzung der großen Labs als strategisch für ihren Erfolg erkannt. Auch neue Technologien wie Cloud-Lösungen oder schneller Zugang zu HPC Ressourcen gehören dazu. Einige Communities haben sich weltweit zum Teil bereits selbst organisiert, um auf Community Ebene eigene Virtuelle Private Netze (VPN) zu schaffen, im Bereich der Teilchenphysik z. B. das LHCONE. Gleiches findet man in der Meteorologie.

Die Digitalisierung in Wissenschaft und Gesellschaft führt zu neuen Arbeitsweisen der Wissenschaftlerinnen und Wissenschaftler, weg von einer ortsgebundenen Bearbeitung von Daten und hin zu vernetzten und über Organisationsgrenzen hinausgehenden Strukturen. Schon heute verlangt die Wissenschaft von einer Informationsinfrastruktur u. a.

- ▶ einen einfachen und möglichst stoßkantenfreien Zugang zu Ressourcen zu haben,
- ▶ flexibel zu sein, um die hohe Dynamik in der Entwicklung von Wissenschaft und Informationstechnologie aufgreifen zu können,
- ▶ Technologien um Daten und Projekte zu finden und effizient auf sie zuzugreifen,
- ▶ Werkzeuge zur Zusammenarbeit bei der Planung, Durchführung und Veröffentlichung wissenschaftlicher Projekte; Analoges gilt auch für die wissenschaftliche Infrastruktur und die Verwaltung,
- ▶ Softwarewerkzeuge und -komponenten, die in einer stabilen Version offen und wiederverwendbar bereitgestellt werden und für die unter anderem sichergestellt ist, dass damit gewonnene Ergebnisse reproduzierbar sind.

Zur Vergrößerung der Anziehungskraft der Arbeitsplätze können auch Arbeitsmodelle beitragen, die jederzeit von praktisch jedem Ort aus Forschung, Datenanalyse, Simulationen etc. möglich machen („services engineered to overcome the constraints of geography“ (ESNET 2015)). Studien zeigen, dass mobiles Arbeiten als sehr positiv für die Vereinbarkeit von Familie und Beruf erlebt wird (BMFSFJ 2016).

Die Helmholtz-Gemeinschaft ist mit ihrer Vielzahl von großen und komplexen Forschungsinfrastrukturen einzigartig. Mit diesen Forschungsinfrastrukturen wird eine große Breite von Forschungsthemen abgedeckt. Die bereitgestellten Instrumente werden durch Wissenschaftlerinnen und Wissenschaftlern aus den Zentren und insbesondere auch anderen wissenschaftlichen Einrichtungen genutzt. Dabei ändert sich die Art der Forschung. Wurde früher eher Community- oder Zentren-spezifisch gearbeitet, so steht heute insbesondere auch durch technische Fortschritte auf den Gebieten der Informationstechnologie Zusammenarbeit und verteilte Kompetenz im Vordergrund. Auch in der Helmholtz-Gemeinschaft spannen sich schon heute viele Themen über mehrere der sechs Forschungsbereiche und werden an verschiedenen Zentren mit unterschiedlichem Schwerpunkt bearbeitet (Helmholtz 2018). Diese Entwicklung wird sich fortsetzen. So werden Großgeräte verstärkt Zentren-übergreifend aufgebaut, betrieben und vermehrt gemeinsam genutzt werden.

Dieses wird nur möglich sein, wenn auch die Anbindung an eine herausragende Informations-Infrastruktur, also die Bereitstellung von Daten-, Computing-, Netzwerk- und Software-Infrastrukturen gegeben ist. Ein prominentes Argument ist der Umgang mit der schnell wachsenden Datenmenge (BIG DATA), siehe auch Abbildung 1, ein Thema, was alle Helmholtz-Zentren betrifft. Große Forschungsinfrastrukturen gehen heute Hand in Hand mit großen Datenmengen und den Herausforderungen, diese schnell zu verarbeiten. Schon heute produziert z. B. die Photonen-Quelle Petra III am DESY 15 PB/Jahr, der High Lumi LHC am CERN, an dem mehrere Helmholtz-Zentren beteiligt sind, wird nach Inbetriebnahme mindestens 1.000 PB/a (1 ExaByte) liefern.

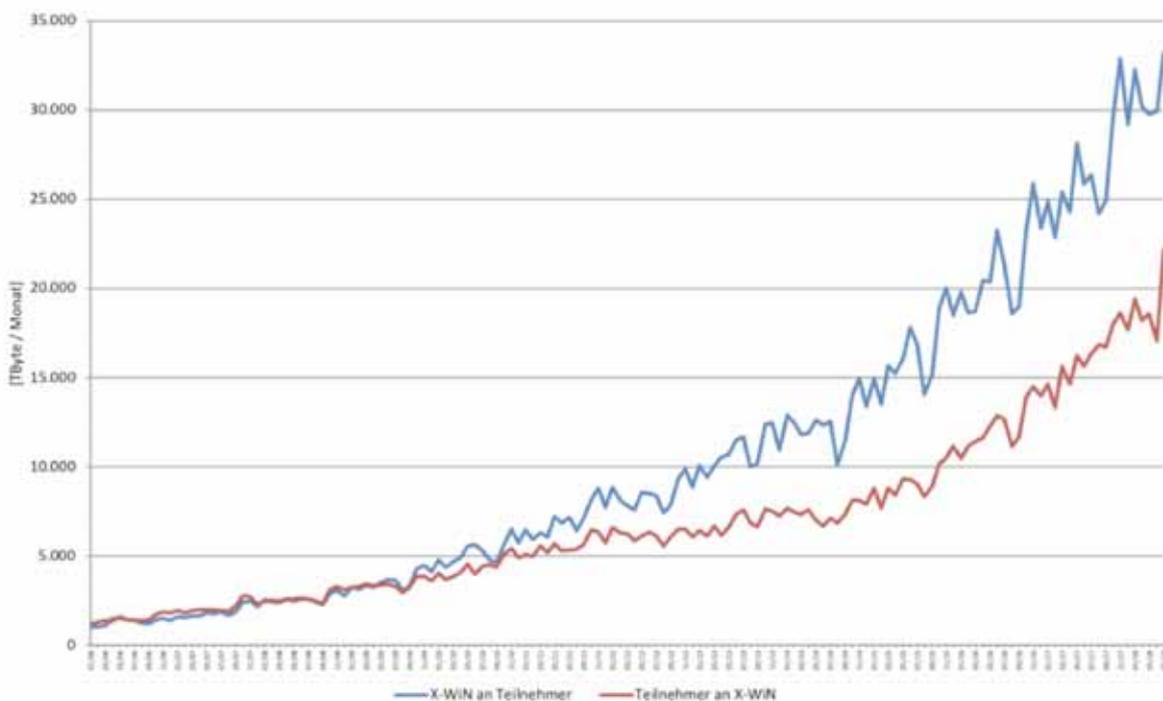


Abbildung 1: Datenvolumen im Forschungs- und Ausbildungsnetzwerk XWiN des Deutschen Forschungsnetzwerks DFN
Anmerkung: Der Vergleich mit der Entwicklung der Anzahl der Nutzer oder der Verschiebung der Nutzerstrukturen im Netz des DFN zeigt nur wenige Veränderungen bei den datenintensiven Nutzern im Betrachtungszeitraum.

Gerade die große Menge originärer Daten, die die Forschung in den Helmholtz-Zentren erarbeitet, wird immer wieder als Alleinstellungsmerkmal bzw. als bedeutender „Datenschatz“ der Helmholtz-Gemeinschaft betont, den es zu nutzen gilt. Eine Besonderheit des Datenverkehrs zwischen den Helmholtz-Zentren besteht in dem Bedarf, in kurzer Zeit große Datenmengen zu transportieren. Im Vergleich zum allgemeinen Internet veranschaulicht das US-NREN ESnet dies mit dem Vergleich von Elefanten, die im Forschungsnetz transportiert werden müssen, und vielen Mäusen, die im allgemeinen Internet transportiert werden müssen (Bell 2015). Aber auch kleinere Datenmengen, oft mit verteiltem Speicherort, können komplexe Strukturen besitzen, die eine umfangreiche und langwierige Analyse erfordern. Wichtig für den Erfolg der Forschung ist daher ein Dienst-Portfolio, das die Breite der Nutzerschaft unterstützt.

Neben der Bereitstellung der notwendigen Hardwarekomponenten spielen insbesondere die bereitgestellten Dienste eine entscheidende Rolle und machen für die einzelne Wissenschaftlerin/ den einzelnen Wissenschaftler in der Nutzung eines Großgerätes einen entscheidenden Unterschied. Auch auf diesem Gebiet steht die Helmholtz-Gemeinschaft in einer Konkurrenz und benötigt Dienste, die eine wissenschaftliche Ernte aus Erkenntnissen, Publikationen, Daten und Software ermöglichen.

Exemplarisch zeigen die folgenden Science Cases diese veränderte Arbeitsweise der Wissenschaft aus drei Forschungsbereichen auf.

SCIENCE CASE FORSCHUNGSBEREICH ERDE UND UMWELT

Beobachtungsdaten spiegeln den aktuellen Zustand des Erdsystems wider und dokumentieren damit im Zeitverlauf die Entwicklungen des Klimas und der Umweltbedingungen von der Vergangenheit zur Gegenwart. Gleichzeitig sind sie die Basis für Modelle, um daraus abgeleitete Aussagen für die Zukunft treffen zu können. Dazu betreiben die Helmholtz-Zentren in der Ozean-, Polar- und Küstenforschung diverse Forschungsplattformen, die einmalige Beobachtungsdaten in Echt- oder quasi-Echtzeit liefern. Dies erfolgt durch vielfältige und leistungsfähige Sensoren mit zunehmender räumlicher, zeitlicher und spektraler Auflösung, woraus immer größere Datenmengen entstehen.



Quelle: Alfred-Wegener-Institute/FRAM/Sabine Lüdeling

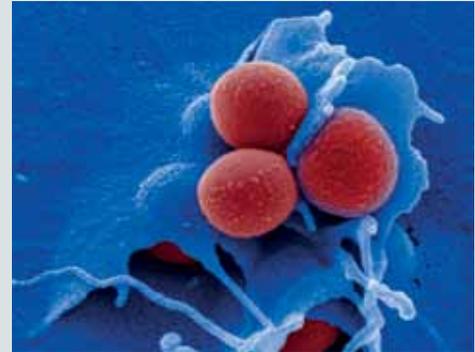
Das in der Polarforschung bewährte Konzept der zentralen Speicherung von Beobachtungsdaten soll in Zukunft für alle Schiffs- und andere Sensorplattformen der Ozeanforschung angeboten werden. Dies ersetzt nicht nur die bisherige Praxis dezentraler, durch die Nutzenden verwalteter und inhärent unsicherer „Mikro-speicher“ (externe USB-Platten, NAS-Speicher), sondern macht diese Daten direkt einer hoch performanten Analytikplattform zugänglich. Somit wird auch erstmals ein sensor-, institutions- bis hin zu disziplin-übergreifender Zugriff zur (effizienten) Aggregation einfach möglich sein. Zugleich wird der zuverlässige Datentransfer vom Sensor zur „Landstation“ technisch klar geregelt und überwachbar. Diese Beobachtungsdaten-Infrastruktur – mit Unterstützungs-, Überwachungs- und Rückwirkungs-Möglichkeiten auf Sensorik und Plattformen auf der einen und Analytik-Möglichkeiten auf der anderen Seite – besteht aus vielen Einzelkomponenten und Diensten, die auf einer (förderierten) Cloud-Plattform unter Mitwirkung aller beteiligten Zentren entwickelt, implementiert, angeboten und unterstützt werden. Der Wechsel zu einer derartigen aktiven Datenhaltung ermöglicht darüber hinaus eine Vielzahl bisher nicht zu realisierender oder nicht skalierbarer Services, etwa

- ▶ Near Real Time Decision Support auf der Basis verteilt gesammelter, disziplinübergreifender Daten (siehe auch MOSES (2018) Usecase „Elbe-Flut“),
- ▶ regelmäßiger Abgleich mit internationalen Datensammlungen, insbesondere fest zugesagte Lieferungen an globale Datenservices (siehe auch GEOSS (2018)),
- ▶ eigene Cloud-Services zur Bedienung von Datenerfassungs-„Apps“ auf Mobilgeräten aller Art (ohne dass deren Daten dann von Herstellern „gefangen gehalten“ werden),
- ▶ gemanagte Kopien einmaliger Daten „off-site“ (z. B. an einem anderen Helmholtz-Zentrum, unter anderer Technologie und Administration), die aus Sicherheits- oder Verfügbarkeitsgründen notwendig sind,
- ▶ das Vorhalten *einer* Kopie von Referenzdatensätzen für alle Nutzer (Sentinel-Daten und -Produkte),
- ▶ Herstellen von Sammlungen zur offline-Verwendung an Bord von Schiffen oder anderen Forschungsplattformen oder -standorten mit geringer Kommunikationsbandbreite, u.v.m.

Für die unterschiedlichen Stationen im Lifecycle der Daten müssen auch die entsprechenden Software Komponenten bereitstehen. Aufnahme- und Auswertesoftware wird ständig modifiziert und den Experimentbedingungen und Analysezielen angepasst. Die verteilt arbeitenden Wissenschaftler und Techniker benötigen dafür geeignete Methoden und Plattformen für ein modernes, agiles und nachvollziehbares Software-Engineering bis hin zur Publikation der Softwarekomponenten mit den Daten.

SCIENCE CASE FORSCHUNGSBEREICH GESUNDHEIT

Moderne Forschung in den Lebenswissenschaften setzt immer mehr auf Hochdurchsatz-Technologie. Diese sogenannten 'Omics'-Technologien erzeugen Daten in erheblichem Umfang, die zunächst gesammelt und vorverarbeitet werden müssen. Dabei kommen Bioinformatik-Pipelines zum Einsatz, die in ihren Modulen durchaus selbst auch Gegenstand von Forschungsarbeit sind. Nach der Prozessierung gehen die Daten in lebenswissenschaftliche Analysen ein, bei welchen in zunehmendem Umfang auch Machine Learning Methoden auf unterschiedlichen Skalen zum Einsatz kommen bis hin zu Deep Learning auf HPC-Systemen.



Quelle: Helmholtz-Zentrum für Infektionsforschung GmbH

'Omics'-Technologien werden an allen Forschungszentren des Forschungsbereichs Gesundheit betrieben. Ein besonderer Fall ist die Humangenomsequenzierung auf der HiSeq X Ten-Plattform in Heidelberg. Diese Plattform stellt eigene Anforderungen an die Informationstechnologie, die lokal in Heidelberg bedient werden. Um die vorhandenen Technologien bestmöglich zu nutzen, wird innerhalb des Forschungsbereichs eine Öffnung der Omics-Plattformen für Projekte anderer Zentren angestrebt. Daraus leiten sich unmittelbar die zu lösenden Fragen nach Identitäts- und Accessmanagement und Ressourcenallokation für externe Kooperationspartner ab. Ferner geht damit zwangsläufig ein Anstieg des internen Datenaustausches einher.

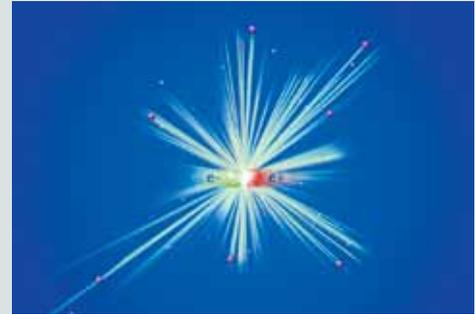
Aus Gründen der Effizienz und des Datenschutzes besteht neben dem Bedarf an Datenaustausch auch das Ziel, ausführbaren Code über Container auszutauschen. Dies erlaubt es unter anderem, einzelne Module in Standard-Pipelines der bioinformatischen Vorverarbeitung durch eigene Software zu ersetzen. Um einen solchen Austausch von ausführbarem Code zu ermöglichen, ist jedoch ein Mindestmaß an Qualitätssicherung bei der Softwareentwicklung genauso Voraussetzung wie entsprechend aufgesetzte Rechteschemata sowie Zugriffsmöglichkeiten auf dezentrale Computing-Services.

Ein weiterer Aspekt der Gesundheitsforschung ist in diesem Zusammenhang der Umgang mit Kohorten- und Klinik-Daten. Im Umgang mit sensiblen Daten ist es besser, die Algorithmen zu den Daten zu bringen. So kann beispielsweise ein Projektpartner eines Konsortiums durchaus berechtigt sein, mit gewissen Daten zu arbeiten, ist aber gleichzeitig nicht berechtigt, die Daten an seiner Heimatinstitution abzulegen. Der Einsatz von Cloud-Technologie erlaubt es, dieser besonderen Situation Rechnung zu tragen. Der strukturierte Umgang mit Zugriffsrechten und Identitäten über Organisationsgrenzen hinaus ist vor dem Hintergrund rechtlicher Rahmenbedingungen im Umgang mit sensiblen Daten von herausgehobener Bedeutung.

Wie in den anderen Forschungsfeldern auch ist die technologische Entwicklung in den Lebenswissenschaften ebenso durch eine enorm hohe Dynamik charakterisiert. Insbesondere die 'Omics'-Technologien haben sich im vergangenen Jahrzehnt zum massiven Treiber von Wissenschaft entwickelt. Eine ICT-Infrastruktur, welche in der Lage ist, diese hohe Dynamik abzubilden, indem traditionelle Grenzen und Barrieren zwischen Organisationseinheiten überwunden werden, sodass ein vereinheitlichter, virtueller Forschungsraum entsteht, wird damit zu einem entscheidenden strategischen Faktor, um an der Weltspitze führend bleiben zu können. ICT muss in die Lage versetzt werden, Enabler von organisationsübergreifender Forschung zu werden und nicht limitierender Faktor. Deshalb bilden die in diesem Antrag adressierten Schwerpunkte eine Grundvoraussetzung, um im Bereich der Multi-omics-Technologien bestmöglich aufgestellt zu sein.

SCIENCE CASE FORSCHUNGSBEREICH MATERIE

Der Forschungsbereich Materie erforscht grundlegende Fragen zu den Bestandteilen der Materie und den zwischen ihnen wirkenden Kräften. Neben den wichtigen Erkenntnissen zum grundlegenden Verständnis von Materie ermöglichen die Forschungen vielfältige Erkenntnisse in verschiedenen Wissenschaftsfeldern. So eröffnen Röntgenlaser einen neuen Blick auf Alzheimer-Proteine, Reihenuntersuchungen von Enzymprozessen werden automatisiert, intakte Viren werden entschlüsselt und die Wirksamkeit von Katalysatoren verbessert.



Quelle: Deutsche Elektronen-Synchrotron DESY

Ein wesentlicher Faktor um, solche Erkenntnisse zu gewinnen, ist die Zusammenarbeit der Helmholtz Forschungszentren untereinander, die in den letzten Jahren stark zugenommen hat. So werden große Forschungsinfrastrukturen wie z. B. PETRA III oder das European XFEL von mehreren Helmholtz-Zentren gemeinsam mit internationalen Partnern betrieben.

Ein anderes Beispiel ist HIBEF, ein gemeinsam von DESY, dem Helmholtz-Institut Jena und dem HZDR getragenes Projekt um ultra-intensive Kurz- und energiereiche Langpuls-Lasersysteme sowie gepulste Hochfeld-Magnetspulen am European XFEL zur Verfügung zu stellen. Aufgrund der hohen zeitlichen Auflösung werden hier sehr große Datenmengen erzeugt und möglichst in Echtzeit analysiert.

Der gemeinsame Aufbau und Betrieb der Großgeräte durch räumlich verteilte Zentren macht nicht nur zwingend eine hervorragende Vernetzung erforderlich, sondern erfordert auch die Bereitstellung kollaborativer Services und gemeinsamer Entwicklungsumgebungen. Die Anforderungen dafür sind u. a.

- ▶ HPC- und Speicher- Ressourcen für Simulation und Auswertung stehen ortsunabhängig zur Verfügung,
- ▶ große Datenmengen sind schnell und sicher transferierbar und werden wenn nötig verteilt bearbeitet und archiviert, dazu sind Qualitätsanforderungen an die Services zu erfüllen, die derzeit nicht gesichert sind (u. a. gesicherte Bandbreite, Latenz, IT-Sicherheit),
- ▶ die gesamte Kommunikation, das Projekt- und Dokumentenmanagement für die Teams sind von jedem Ort ohne Einschränkungen (unter Beachtung von Sicherheit und Datenschutz) nutzbar,
- ▶ Entwicklungen (vor allem für die komplexe Software) verlaufen von Beginn an in gemeinsamen Entwicklungs- und Testplattformen mit definierten Prozessen und Qualitätsmanagement.

Komplexe gemeinsame Projekte im Fachbereich Materie, wie hier am Beispiel HiBEF erläutert, können heute ohne Zentren-übergreifende cutting edge Vernetzung, kollaborative Plattformen und moderne, agile Entwicklungsmethoden und -Systeme, vor allem für die Software, nicht erfolgreich realisiert werden.

Die Helmholtz-Gemeinschaft sollte vor diesem Hintergrund die bisher Zentren-fokussierte ICT-Versorgung schrittweise hin zu föderierten Helmholtz-weiten (und ggf. darüber hinaus gehenden) Lösungen entwickeln, die das Potenzial moderner ICT-Lösungen organisationsübergreifend zugänglich machen.

Diese Notwendigkeiten werden im vorgelegten Konzept aufgegriffen, in dem eine Vision einer Service Infrastruktur entwickelt wird, die von der physikalischen Vernetzung der Zentren mit gemeinsamen Autorisierungsdiensten über Cloud-Services für die wissenschaftliche Nutzung bis zu modernen, agilen Methoden und Plattformen für den Umgang mit wissenschaftlicher Software geht.

Während sich andere im Rahmen des Helmholtz-Inkubators entwickelte Plattformen den wissenschaftlichen Methoden und Analyseverfahren widmen, konzentriert sich dieser Antrag auf die Dienste und die Infrastrukturen, die jede Wissenschaftlerin und jeder Wissenschaftler als gegeben voraussetzt und die für eine moderne Spitzenforschung unerlässlich sind. Notwendige Services für eine ICT-Infrastruktur und wie diese von den wissenschaftlichen Nutzern und Partnern gesehen werden, stellt das folgende Bild dar. Es veranschaulicht, dass die Sichtbarkeit von den Methoden und Datenservices bis zur Infrastruktur abnimmt. Dabei werden unter den jeweiligen Services exemplarische Nutzer benannt und auf der linken Seite exemplarische Projekte der Helmholtz-Gemeinschaft.

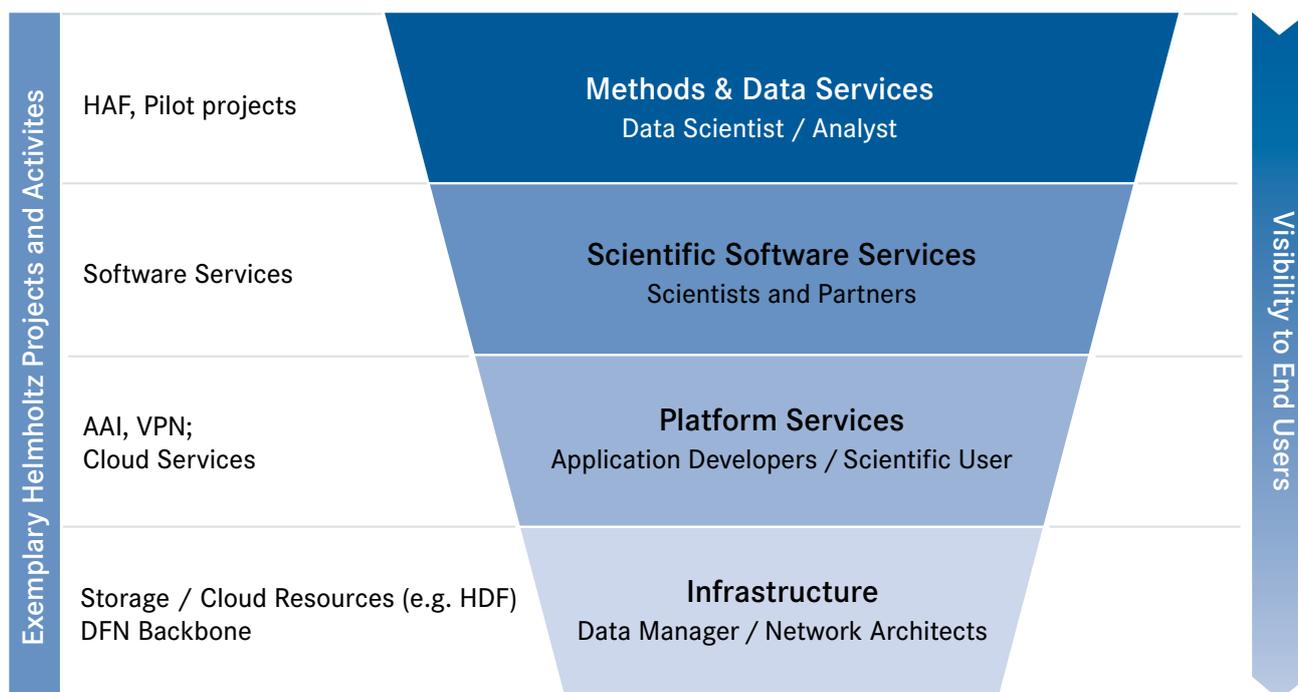


Abbildung 2: Services und ihre Sichtbarkeit

Um Zentrums- und Gemeinschafts-übergreifend deutlich voranzukommen, sind Aktivitäten auf all diesen Service-Ebenen erforderlich, also nicht nur im Bereich Methoden und Daten, sondern auch in den Bereichen Applikations-Plattformen, Infrastrukturen und Software.

3 ZIELE

Das grundlegende Konzept für die hier vorgeschlagene HIFIS-Plattform wurde im September 2017 der Mitgliederversammlung der Helmholtz-Gemeinschaft vorgestellt. Der Plattformvorschlag beinhaltet drei Kernziele:

1. Aufbau einer Helmholtz-weiten Serviceplattform für die Cloud, die an jedem Zentrum und in vielen wissenschaftlichen Projekten genutzt wird und über welche die Helmholtz-Zentren ihre best-of-practice Lösungen und Kompetenzen anderen Zentren und externen Partnern zur Verfügung stellen können (ein Marktplatz für Anwendungen). Diese innovative Cloud-Lösung soll eine Strahlkraft und Skalierbarkeit aufweisen, welche die Helmholtz-Gemeinschaft als large-scale Service-Provider für nationale und internationale wissenschaftliche Aufgaben qualifiziert (wie z. B. NFDI, EOSC u. a.).
2. Die Zentren-übergreifende Nutzung von Services und Ressourcen in kollaborativen Projekten setzt eine hervorragende Vernetzung der Zentren voraus, die ebenfalls in diesem Konzept spezifiziert wird. Ziel ist die Realisierung von Services für ein hoch performantes Helmholtz Backbone, eines (logischen) Datenaustausch- und Kommunikationsnetzes zwischen den Zentren. Dieses Netzwerk soll auf existierende Infrastrukturen aufsetzen z. B. dem Deutschen Forschungsnetz (DFN) und einen logischen Verbund mit definierten Qualitätsmerkmalen wie z. B. definierter Bandbreite, Sicherheit und einheitlichen Zugriffsmechanismen liefern. Basisdienste wie z. B. eine Authentifizierungs- und Autorisierungs-Infrastruktur (AAI) sollen bereitgestellt werden.
3. Software ist nicht nur eine Grundlage des modernen Forschungsprozesses und die Basis für den langfristigen, stabilen Betrieb der Forschungsinfrastrukturen, sondern auch selbst ein wichtiges Ergebnis (neben Publikation und Daten). Ein Ziel des Antrags ist es, die Erstellung wissenschaftlich-technischer Software der Helmholtz-Gemeinschaft als Ergebnis eines professionellen Entwicklungsprozesses mit hoher Qualität zu unterstützen und dabei einen nachhaltigen Umgang mit der Software zum Zwecke der Nachvollziehbarkeit von Ergebnissen und der möglichen breiten Nutzung zu gewährleisten. Dafür sollen die Rahmenbedingungen auf allen Ebenen, beginnend mit der Ausbildung über die technische und organisatorische Unterstützung im Entwicklungsprozess bis hin zur Publikation und der Pflege der Softwareprodukte geschaffen werden.

Die Anbindung dieses Vorhabens an existierende und geplante Projekte in den Communities (z. B. MOSES, ESM, Digital Earth, MOSAIC) ist von großer Bedeutung. Hierbei wird stark auf laufende Entwicklungen in anderen Projekten und auf bereits existierende Lösungen zurückgegriffen. Die Plattform HIFIS kann auch ganz wesentlich die weiteren Pilotprojekte und Plattformen des Inkubators unterstützen, die einerseits auf den infrastrukturellen Services aufsetzen und andererseits die Bereitstellungs-Services für neue Anwendungen nutzen können (siehe hierzu auch die Ausführungen zu den Schnittstellen zu den anderen Inkubator-Plattformen in Kapitel 9.6).

In der Helmholtz Data Federation (HDF) wird bisher vorwiegend Hardware an sechs Zentren mit dem Schwerpunkt der Unterstützung besonders großer wissenschaftlicher Vorhaben gefördert. Bei dem vorliegenden HIFIS-Antrag liegt der Schwerpunkt auf Services und Software, die Helmholtz-weit bereitgestellt werden. Insofern greift dieser Antrag Entwicklungen in der HDF auf und fügt diese mit Services auf anderen Technologieebenen zusammen. Während HDF und Helmholtz Analytics Framework (HAF) in der jetzigen Ausbaustufe zumeist wichtige große Community-Projekte mehrerer Forschungsbereiche mit spezifischen Lösungen der teilnehmenden Zentren und Partnern adressieren, zielt das Konzept des HIFIS-Antrages auf Services, die auch in der Fläche und für kleinere Projekte genutzt werden können, bezieht aber dabei die Helmholtz Data

Federation mit ein. Schnittstellen sind dabei z. B. spezialisierte Infrastrukturelemente und Basisdienste wie das zentrale Identity- und Access Management (IAM). Darüber hinaus sind ebenfalls Schnittstellen zu Projekten zu berücksichtigen, die über die Helmholtz-Gemeinschaft hinausgehen. Beispiele sind hier u. a. Generic Research Data Infrastructure (GeRDI) und European Open Science Cloud (EOSC).

In der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen (Allianz 2018) sind die in diesem Antrag adressierten Themen ein wesentlicher Bestandteil, der über die Deutsche Forschungsgesellschaft (DFG) auch in neue Förderprojekte führen wird. So sind die beiden Themen „Föderierte IT-Basisinfrastruktur“ und „Digitale Werkzeuge und Dienste“ wichtige Arbeitsgebiete der neuen Arbeitsperiode bis 2022. Diese erfordern es, umgehend praxiserprobte Lösungsangebote einzubringen, so wie sich die Fraunhofer Gesellschaft (FhG) und die Max Planck Gesellschaft (MPG) als Provider mit zentralen personellen Ressourcen organisiert haben und bereits entsprechende Aktivitäten vorantreiben.

Auf nationaler und internationaler Ebene soll dieser Vorschlag gerade auch im Verbund mit den wichtigen Ressourcen, die über die HDF bereitgestellt werden, die Bedeutung und Sichtbarkeit der Helmholtz-Gemeinschaft und auch den Einfluss weiter steigern. Andere Forschungsorganisationen in europäische Länder positionieren sich deutlich in diesem Sinne, hier sei z. B. auf das STFC in UK verwiesen, siehe Kapitel 9.4. Gerade beim Aufbau der Nationalen Forschungsdateninfrastruktur (NFDI) und in der nationalen Zusammenarbeit der außeruniversitären Forschungsorganisationen ist die starke Mitwirkung der Helmholtz-Gemeinschaft unabdingbar und erfordert eine Zentren-übergreifende Organisation.

4 STRUKTUR DER AKTIVITÄTEN

Die Plattform „Helmholtz Infrastructure for Federated ICT Services“ (HIFIS) besteht aus den Teilkomponenten **Cloud Services**, **Backbone Services** und **Software Services**, die eng miteinander verknüpft sind, jedoch jeweils eigene Kompetenzgebiete darstellen. Mit HIFIS werden Strukturen geschaffen, die gemeinschaftsweite Relevanz haben.

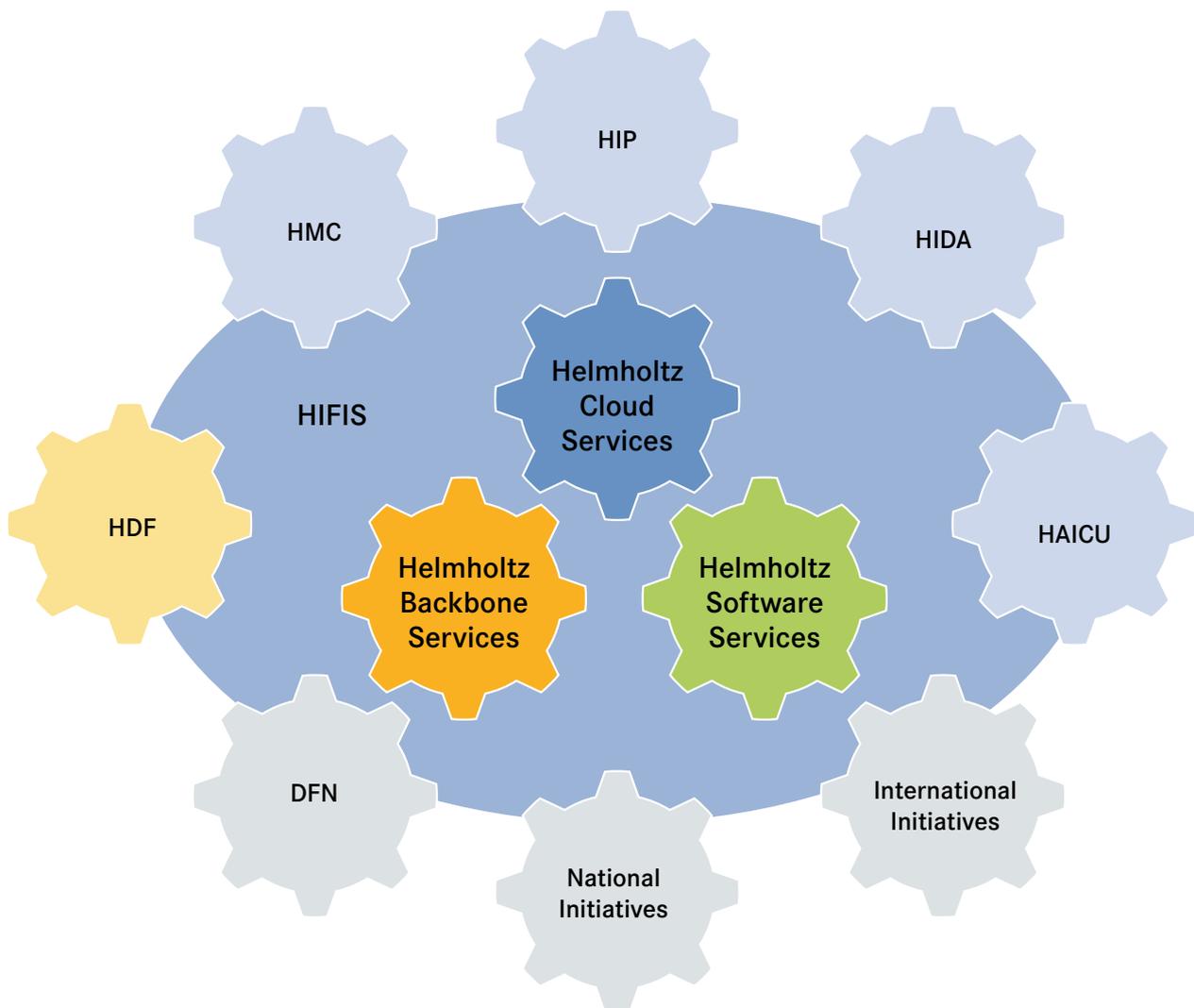


Abbildung 3: HIFIS Services und deren Vernetzung

Spezifisches Anliegen der Komponente „**Cloud Services**“ ist die Bereitstellung einer föderierten Cloud-Plattform für Services, die der gesamten wissenschaftlichen Community und den Partnern in der Fläche (also nicht nur in ausgewählten Großprojekten) angeboten werden. Die Föderation setzt sich vorrangig aus existierenden Lösungen der verschiedenen Anbieter zusammen und legt eine „Metaebene“ über diese, um den Nutzern einen einheitlichen Zugang zu ermöglichen. Diese wird dann schrittweise um innovative Lösungen erweitert.

Durch die Komponente „**Backbone Services**“ erfolgt die Bereitstellung einer hoch performanten Netzwerkinfrastruktur zwischen allen Zentren. Die Umsetzung der notwendigen Infrastruktur in den einzelnen Zentren ist nicht Gegenstand der HIFIS-Aktivitäten, sondern wird durch jedes einzelne Zentrum selbst durchgeführt. Die

Netzwerkinfrastruktur soll insbesondere die Zentren-übergreifende Nutzung wissenschaftlicher Instrumente ermöglichen sowie eine Reihe von Basis-Services, die für kollaboratives Arbeiten notwendig sind, bereitstellen. Das physikalische Netz zusammen mit den Basisdiensten bilden die Voraussetzung und das Fundament für eine Vielzahl höherwertiger Dienste wie z. B. Cloud-Services.

Über die Komponente „**Software Services**“ werden zum einen Forscherinnen und Forscher ausgebildet und unterstützt, um sie zur Entwicklung und Publikation von qualitativ angemessener, nachhaltiger Software zu befähigen. Zum anderen wird die Nachnutzung vorhandener Software durch Forscherinnen und Forscher befördert, indem Werkzeuge zum Finden und Nutzen von geeigneten Software-Lösungen bereitgestellt und Ansprechpartner bzw. Know-how-Träger vermittelt werden sowie Unterstützung beim Aufbau von Communities für konkrete Forschungssoftware gewährt wird.

Eine enge **Verzahnung der HIFIS-Plattform mit den Aktivitäten der HDF** wird sichergestellt. Schnittstellen werden sich nach der bisherigen Planung vor allem in den Bereichen AAI/IAM und Cloud Services ergeben. So bearbeitet die HDF beispielsweise in ihrem Arbeitspaket 1 „Authentication and Authorisation Infrastructure (AAI)“ die Thematik des Zugangs von föderierten Nutzern zu föderierten Diensten mit dem lokalen Konto der Benutzer. Arbeitspaket 6 „Föderierte Cloud“ befasst sich mit dem entsprechenden Zugang zu HDF-Daten und -Diensten. Insbesondere beim Thema AAI/IAM können die Arbeiten im HDF-AP1, welches wiederum auf Arbeiten in europäischen Projekten fußt, als ideale Grundlage für eine Helmholtz-weite AAI/IAM-Lösung mit voller Kompatibilität zur nationalen DFN-AAI und zu europäischen AAI/IAM-Lösungen im Kontext GEANT/EduGain und EOSC dienen.

In der Folge der Abstimmung über Arbeitsinhalte zwischen HDF und HIFIS im Verlauf der HIFIS-Antragsentwicklung wurde festgelegt, dass die Arbeiten des HDF-AP8 (Helmholtz-VPN-Ring) bis auf einen Prototyp am AWI an HIFIS ausgelagert werden könnten.

Die gewünschte enge Verzahnung wird auf der Arbeitsebene auch über die sowohl an HDF als auch an HIFIS beteiligten Personen gesichert. Weiterhin gibt es engste personelle Überschneidungen auf der steering-Ebene zwischen HDF und HIFIS, insbesondere durch Volker Gülzow (DESY) als einen der HIFIS-Berichterstatter und gleichzeitiges Mitglied des HDF Executive Board (EB).

Über diese Vorgehensweise bzw. Verzahnung ist sichergestellt, dass vorhandenes Know-how in den Zentren genutzt wird, an die bereits geleisteten Arbeiten angeknüpft und diese effizient ergänzt werden sowie Interoperabilität hergestellt und Doppelentwicklungen vermieden werden.

Weiterhin erfolgen im Rahmen der dargestellten Arbeitspakete Gespräche mit weiteren externe Kompetenzträgern und Initiativen an Helmholtz-Zentren (wie z. B. dem Team „Scientific IT Systems“ – einer gemeinsamen Einrichtung des Peter Grünberg Instituts (PGI) und des Jülicher Zentrum für Forschung mit Neutronen (JCNS)).

5 ARBEITSPROGRAMM

5.1 CLOUD SERVICES

STATUS QUO

Cloud-Dienste finden zunehmend Verwendung in der Wissenschaft. Eine abgrenzende Definition von Cloud-Diensten ist nicht vorhanden (und wahrscheinlich auch nicht sinnvoll möglich). Das National Institute of Standards and Technology (NIST) beschreibt Cloud Computing als ein Modell für den allgegenwärtigen, bequemen und bedarfsgerechten Netzwerkzugriff auf einen gemeinsamen Pool von konfigurierbaren Computerressourcen (z. B. Netzwerke, Server, Speicher, Anwendungen und Dienste), die mit minimalem Verwaltungsaufwand oder Service Provider-Interaktion schnell bereitgestellt und freigegeben werden können. Es nennt fünf charakteristische Eigenschaften, die Cloud-Dienste ausmachen (on-demand self-service, broad network access, resource pooling, rapid elasticity, measured service) sowie drei Arten von Modellen zur Dienstleistung: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) und Software as a Service (SaaS).

Die europäische Netzwerkorganisation GEANT¹ hat in der Umfrage in den nationalen Forschungs- und Ausbildungsnetzwerken (NREN) für das Jahr 2016 erhoben, dass in 50 % bis 60 % der von GEANT befragten NREN Cloud Services angeboten werden. „Cloud-Storage ist der häufigste Service von NRENs, 62 % liefern derzeit oder planen die Bereitstellung von Cloud-Storage in den nächsten 12 Monaten. Infrastructure as a Service (General-Purpose Cloud Computing Services) wird von 61 % der NRENs angeboten, was darauf hindeutet, dass diese beiden Servicetypen für den F&E-Sektor von entscheidender Bedeutung sind und bleiben. Software as a Service und Platform as a Service ergänzen das normale Leistungsportfolio der meisten NRENs.“ Die meisten NRENs folgen einem „Build“-Modell und nicht einem „Buy“-Modell für Commodity-Dienste. Eine Ausnahme stellt Software as a Service dar, den NRENs kaufen. Dies ist wahrscheinlich auf das breite Spektrum der benötigten Softwarelösungen und die fehlenden Fähigkeiten zurückzuführen, um eine solche Vielfalt kostengünstig zu unterstützen (GEANT-Compendium 2016).

Auf europäischer Ebene werden eine Reihe von Cloud-Entwicklungen vorangetrieben, die z. B. in der European Open Science Cloud mit dem Open Science Ansatz verbunden werden (siehe z. B. EC (2018)). Ein europäisches Cloud-Modell mit förderierter e-Infrastruktur wird beispielsweise durch EGI erarbeitet bzw. angeboten².

Cloud Services sind auch mit regionalem Zuschnitt realisiert. Ein Beispiel ist **Sciebo**, ein gemeinsames Projekt von Sync & Share NRW, einem Zusammenschluss von 26 Hochschulen und Forschungseinrichtungen in NRW. Sciebo ist ein nichtkommerzieller Cloud-Speicherdienst für Forschung, Studium und Lehre³. Ein weiteres Beispiel ist **bwCloud SCOPE Science, Operations and Education**, dessen Aufgaben der Aufbau und der Betrieb einer standortübergreifenden Infrastruktur zur Bereitstellung von Compute-Ressourcen (virtuelle Maschinen) in Baden-Württemberg sind.

Ein weiterer Anspruch von Cloud Diensten ist die Öffnung und Vernetzung auf der Ebene von entsprechenden Angeboten der Wissenschaftsorganisationen, wie z. B. der Max-Planck Gesellschaft (Yahyapour, R.; Wieder, P. (2013)) und Schwerpunkten in spezifischen Wissenschaftsbereichen wie Cloud Services für Anwender in der Life

1 siehe <https://www.geant.org/About>

2 siehe z. B. <https://www.egi.eu>

3 siehe z. B. <https://www.sciebo.de/projekt/>

Science Forschung und Biomedizin⁴. Ein Beispiel ist die Embassy Cloud, ein IaaS, das EMBL-EBI-Mitarbeitern und externen Organisationen Zugriff auf EMBL-EBI-Datensätze gibt⁵.

ZIEL

Cloud-Dienste eignen sich hervorragend für die gemeinsame, Zentren- und Community-übergreifende Bearbeitung von Forschungsaufgaben und für eine breite Nachnutzung. Wichtig ist dafür, dass sie skalierbar sind, die technischen bzw. rechtlichen Voraussetzungen für die Nutzung gegeben sind sowie ein einfacher und performanter Zugang gewährleistet wird. Gleichzeitig erfordert die Entscheidung für die Nutzung solcher Dienste, dass wichtige Kriterien (von der Servicequalität über IT-Sicherheitskriterien und Datenschutz bis hin zu Verfahrensweisen bei der Beendigung der Nutzung eines Cloud-Dienstes) berücksichtigt und bewertet werden. Die Nutzer werden einer föderierten Cloud-Plattform und den darin angebotenen Diensten nur dann vertrauen (können) und diese nutzen, wenn Servicedesign und Servicequalität transparent sind und die angebotene Betriebsstabilität den gesellschaftlichen Anforderungen der heutigen Zeit an Mobilität sowie orts- und zeitunabhängiger Nutzung gerecht wird.

Spezifisches Anliegen dieser Komponente ist die Bereitstellung einer föderierten Cloud-Plattform für Services, die der gesamten wissenschaftlichen Community und den Partnern in der Fläche (long tail of science) – also nicht nur ausgewählten Großprojekten – angeboten werden. Sie ergänzt damit die z. B. in der HDF oder in Pilotprojekten geschaffenen föderierten Unterstützungsleistungen für Großprojekte sowohl in der Breite (Unterstützung für Projekte aller Größenordnungen) als auch in der Tiefe (Fokussierung auf die Software- und Serviceebene) und stellt anderen Plattformthemen eine Zugriffsebene für die dort entstehenden Services bereit.

Die Föderation setzt sich vorrangig aus existierenden heterogenen Lösungen der verschiedenen Zentren (Anbieter) zusammen und legt, wo möglich und sinnvoll, eine „Metaebene“ über diese, z. B. einen gemeinsamen Zugang. Priorität hat dabei das „Buy“-Modell (im Sinne von Wiederverwendung bereits vorhandener Services), eigene Entwicklungen sind nur ausnahmsweise für Lösungen vorgesehen, die weder in der Helmholtz-Gemeinschaft noch am IT-Markt verfügbar sind und gleichzeitig dringend für die wissenschaftliche Arbeit benötigt werden. Die Föderation wird schrittweise um innovative Lösungen erweitert.

Um die Nutzung von Diensten unterschiedlicher Anbieter für die Zentren und die Forscherinnen und Forscher zu vereinfachen und mit Hilfe von Guidelines zu gewährleisten, dass die für die Helmholtz-Zentren geltenden Rahmenbedingungen eingehalten werden, sind organisatorische und rechtliche Voraussetzungen zu schaffen. Durch transparente Kriterien für Cloud Services wird verhindert, dass innerhalb der Helmholtz-Cloud ein ungehinderter Wildwuchs zu einer unüberschaubaren Zahl von Diensten führt, und erreicht, dass ein in sich konsistentes und geordnetes Serviceportfolio dem Nutzer zur Verfügung steht.

Eine weitere Voraussetzung für die Nutzung von Cloud-Diensten aus einer Föderation ist die Etablierung eines einheitlichen Identitäts- und Access-Managements, das es erlaubt, Dienste einer Organisation (Zentren, Communities, Partner) auch für Mitglieder einer anderen Organisation zugänglich zu machen.

Entwickelt und bereitgestellt werden vor allem Technologien zum Zugang, der Harmonisierung, der Integration und Provisionierung der in den Cloud-Verbund eingebrachten Services. Die Komponente Helmholtz Cloud

4 siehe z. B. de.NBI CLOUD COMPUTING FOR LIFE SCIENCES http://www.denbi.de/localmedia/documents/deNBI_Cloudflyer.pdf

5 siehe z. B. <http://www.embassycloud.org/>

Services schafft somit den organisatorischen und technischen Rahmen, um über das Prinzip der Selbstorganisation ein breit gefasstes Dienste-Portfolio allen Nutzern der Helmholtz-Gemeinschaft und darüber hinaus zu eröffnen. Dabei wird die Anbindung an internationale Entwicklungen (z. B. EOSC auf europäischer Ebene) als essentiell angesehen und es müssen insbesondere die „FAIR“-Prinzipien berücksichtigt werden.

SCHRITTE ZUR ZIELERREICHUNG

Die Entwicklung und Maintenance der Helmholtz Cloud Services erfolgt federführend durch das Helmholtz Cloud Competence Center (CCC). Die Unterarbeitspakete bauen dabei teilweise aufeinander auf, sodass die Abfolge in der zeitlichen Planung der Arbeitspakete berücksichtigt werden muss.

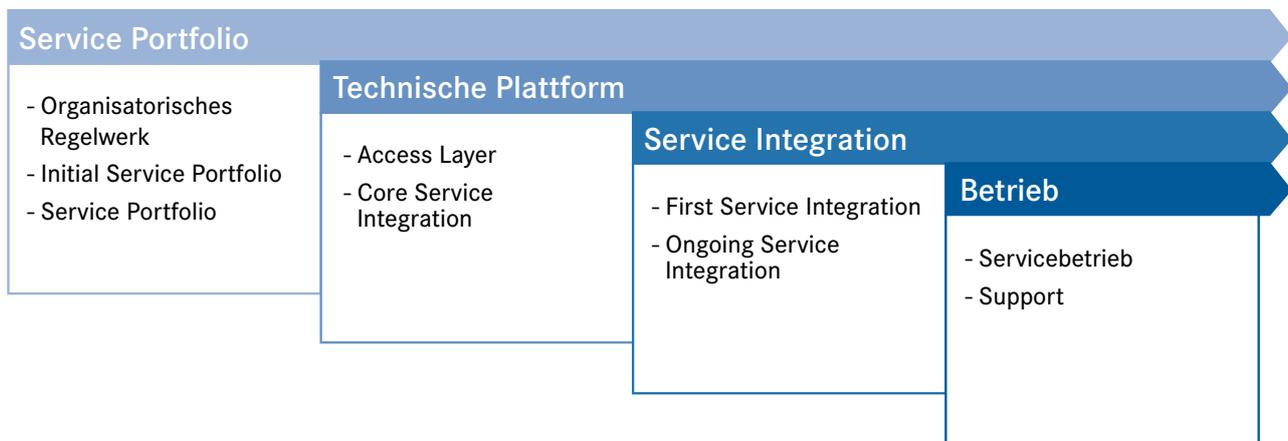


Abbildung 4: Schritte zur Zielerreichung

Am Anfang stehen die organisatorischen Voraussetzungen und die Gestaltung eines **initialen Serviceportfolios** sowie die Definition und Dokumentation der Verfahrensweisen zur stetigen Anpassung des Portfolios. In einem organisatorischen Regelwerk für Cloud-Services wird u. a. beschrieben,

- ▶ wie Service Provider Dienste für die Aufnahme in die Cloud anmelden können, nach welchen Kriterien über die Aufnahme entschieden wird, welche Regelungen gelten sollen, wenn die Bereitstellung eines Services beendet werden soll,
- ▶ wie der gesamte Service Lifecycle für die Cloud Plattform (Service Design, Service Transition, Service Operation) abgebildet wird und
- ▶ wie betriebswirtschaftliche Rahmenbedingungen für Nutzer und Provider gestaltet sind.

Parallel zur Entwicklung des Regelwerkes ist ein initiales Serviceportfolio zu bestimmen, um von Anfang an für eine Attraktivität der Plattform und damit für eine intensive Nutzung zu werben. Wichtigster Baustein dabei ist der Abgleich der von Helmholtz -Zentren angebotenen Dienste für eine Cloud mit den von den Nutzern signalisierten Bedarfen. Beide Aspekte sind in intensiven Bestandsaufnahmen bei den Zentren zu ermitteln. Hervorzuheben ist dabei, dass in dieser Phase auch die direkte Einbindung der HDF und der dort vorhandenen IT-Lösungen geplant ist, um auf diesen aufsetzend die Serviceebene für Nutzer adressierbar zu machen. Bereits identifizierte Services bilden dabei eine Basis für die Ergänzung durch weitere benötigte Dienste. Diese werden in einer gap-Analyse identifiziert und es werden unterschiedliche Providerszenarien geprüft, von der Beauftragung eines Zentrums für die Entwicklung und Bereitstellung eines nachgefragten Services bis hin zur Prüfung von Angeboten Helmholtz-externer Serviceanbieter. Mit potentiellen Serviceanbietern ist zu verein-

baren, welche Arbeiten notwendig sind, um lokale Lösungen Cloud-fähig zu machen, sich als Service Provider zu profilieren und in der Helmholtz-Gemeinschaft einen wichtigen Beitrag bei der digitalen Transformation der Forschungsarbeit zu leisten.

Das Helmholtz Cloud Serviceportfolio ist langfristig und stetig weiterzuentwickeln, um stets innovative und moderne Lösungen innerhalb der Cloud anzubieten. Die dafür notwendigen Regelungen werden Teil des organisatorischen Regelwerks. Die Weiterentwicklung und Pflege des Helmholtz-Gemeinschaft Cloud Service Portfolios geht einher mit dem Ausbau des Serviceangebotes, der Validierung vorhandener Services und einem regelmäßigen Review des Portfolios.

Parallel zur Gestaltung des Serviceportfolios ist damit zu beginnen, die **technische Plattform** für die Helmholtz Cloud zu implementieren. Das Einbringen der technischen Aspekte und der unterschiedlichen Perspektiven der Service Provider sowie der Nutzercommunity ist erforderlich, um über den Access-Layer und das Backbone Nutzer und Serviceprovider zu verbinden. Dieses Ziel wird mit dem Aufbau der Access-Infrastruktur, die sowohl den zentralen Nutzerzugang zu Cloud-Services beinhaltet, als auch integrierbar ist in die jeweilige IT-Infrastruktur der nutzenden Zentren, verfolgt. Verbunden ist diese technische Implementierung mit der Entwicklung von Core-Services, also Diensten, die für das Funktionieren der technischen Cloud-Plattform erforderlich sind. Darunter fallen Lösungen z. B.

- ▶ für die einheitliche Authentifizierung der Nutzer an der Access-Infrastruktur, bei der u. a. auf die Kompatibilität zu adäquaten Entwicklungen auf nationaler und europäischer Ebene zu achten ist und
- ▶ für das Monitoring u. a. im Sinne der Nutzung von Cloud-Diensten.

Für die Authentifizierung wird auf eine entsprechende AAI – Lösung zurückgegriffen, deren Entwicklung im Rahmen der Backbone Services vorgesehen ist.

Nachdem die technischen Voraussetzungen vorhanden sind, das organisatorische Regelwerk definiert ist und das initiale Serviceportfolio abgesteckt ist, beginnt die **Integration der Services** in die Cloud-Plattform. In dieser Phase werden Services durch die Anbieter in der Plattform verfügbar gemacht, getestet und nach erteilter Freigabe für die Nutzer freigeschaltet. Die für einen Service jeweils erforderliche Integrationstiefe in die Service Plattform und die IT der jeweils nutzenden Zentren ist dabei von Service zu Service unterschiedlich.

Mit der Freigabe bzw. Freischaltung zur Nutzung ist für alle Komponenten der Übergang in die **Betriebsphase** vorgesehen. Dauerhafte Aufgaben im Betrieb sind neben der Bereitstellung des Dienstes mit den vereinbarten Merkmalen auch die dauerhafte Überwachung der Servicequalität (Monitoring), die Erfassung und ggf. Abrechnung der Nutzung der Services (Accounting) sowie die Unterstützung der Nutzer im Sinne eines technischen Supports.

Cloud Service Portfolio

Die Erarbeitung von Grundlagen für die Cloud-Nutzung ist für Nutzer und Provider gleichermaßen essentiell, da letztlich eine vertraglich fixierte Zusammenarbeit zwischen beiden zustande kommt. Auch wenn dies in der aktuellen Nutzung von verschiedenen Diensten aus der Public Cloud, wie z. B. Google Docs, von den meisten Nutzern nicht mit der notwendigen Aufmerksamkeit zur Kenntnis genommen wird, ist der Gebrauch solcher Dienste mit einer eindeutigen Vereinbarung zwischen Nutzer und Provider verbunden. Im besten Fall wird zwar zur Kenntnis genommen, dass man einer Vereinbarung zustimmt, aber der Inhalt ist für den Einzelnen nur schwer bewertbar.

In einem organisatorischen Regelwerk ist ein Anforderungskatalog an Services und Service Provider zu erarbeiten, wobei:

- ▶ die sicherheitsrelevanten Anforderungen Konsens in der Helmholtz-Gemeinschaft sein sollen,
- ▶ die Anforderungen sich soweit möglich an vorhandenen Standards orientieren sollen,
- ▶ Anforderungen nur dort wo nötig definiert werden und
- ▶ Transparenz für Provider und Nutzer geschaffen wird.

Das Regelwerk wird sich damit der Problematik stellen, nachvollziehbare Ordnungskriterien für das Portfolio zu erlassen und einem Wildwuchs entgegenzuwirken. Zu dieser Aufgabe gehört es, transparente Prozesse zu definieren und zu dokumentieren, die den Lebenszyklus eines Dienstes in der Helmholtz-Cloud beschreibt, von der Aufnahme eines Services in das Portfolio bis zur Beendigung eines Services.

Für das Accounting wird derzeit davon ausgegangen, dass die Service Provider aus der Helmholtz-Gemeinschaft die Services für die Nutzer in einer ersten Phase kostenfrei erbringen und damit keine nutzungsabhängige Abrechnung erfolgt. Allerdings kann für Services ein Mehraufwand beim Provider entstehen, um den Dienst Cloud-fähig zu machen und der breiten Nutzerschaft zur Verfügung zu stellen. Für diese Fälle sind Mittel vorgesehen, die als Sachmittel in die Finanzplanung eingehen.

Eine Abrechnung der Services (Accounting) ist für einen späteren Zeitpunkt prinzipiell möglich und wird spätestens mit der Einbeziehung externer Services auch voraussichtlich erforderlich. Ein in der Aufbauphase etabliertes Monitoring dient zunächst zur Erfolgskontrolle, in der Betriebsphase kann nach Bedarf auch eine Abrechnung erfolgen.

In der Startphase der Entwicklung des Helmholtz Cloud Service Portfolios erfolgt eine initiale Nutzer-Analyse die unter anderem erfasst, welche Cloud Services verfügbar sind und welche gewünscht werden. Cloud Services, die gewünscht aber nicht verfügbar sind, können in diesem Prozess (ggf. auf der Grundlage einer erweiterten Bedarfs- und Machbarkeitsanalyse) auf eine to do Liste gesetzt werden, die durch Marktrecherchen und ggf. Helmholtz-Eigenarbeit abgearbeitet wird.

Die Bereitstellung der Services erfolgt bevorzugt auf der Basis bereits vorhandener Lösungen an den Zentren nach dem EfA (Einer für Alle) -Prinzip. Das Helmholtz Cloud Service Portfolio kann neben den Cloud Services der Helmholtz-Zentren auch Services umfassen, die von Anbietern außerhalb der Helmholtz-Gemeinschaft zur Verfügung gestellt werden.

Bei der Entwicklung eines breiten Dienste-Portfolios können verschiedene Gruppen von Services differenziert werden: Infrastruktur Services, Community Services und Collaboration Services.

Unter sogenannten Infrastruktur Services fallen folgende Beispiele:

- ▶ Durch immer bessere Messgeräte und neuartigen Methoden bei der Beobachtung bzw. Messung wird verfügbarer Datenspeicher zunehmend eine kritische Ressource. Dieser Trend wird dadurch verstärkt, dass für bereichsübergreifende Forschung Datenquellen verschiedener Zentren bzw. Forschungsbereiche für gemeinsame Analysen kombiniert werden. On-demand-Speicher stellt Ressourcen als Service auf Anforderung zur Verfügung, die dafür notwendige Kapazitätsplanung und Technikbeschaffung konzentriert sich auf wenige Zentren (HDF). Diese Speicherressourcen können mit weiteren Services kombiniert werden, z. B. mit computing Ressourcen.

- ▶ Neben Speicher wird für die Analytik insbesondere auf computing Ressourcen zurückgegriffen. Für analytische Software ist die Bereitstellung von virtuellen Systemen auf Basis von OpenStack oder ESX/ESXi als Service ein Mittel, um die bedarfsgerechte Verfügbarkeit solcher Systeme zu erhöhen. Dieser Service ist orchestrierbar mit On-demand-Speicherdiensten.
- ▶ In Verbindung mit analytischer Software in einem ausführbaren Container können virtuelle Server nicht nur als Infrastruktur bereitgestellt werden, sondern schon vorkonfiguriert für die spezielle Verwendung. Gerade in Verbindung mit den erwarteten Ergebnissen aus dem Inkubator Pilotprojekt Helmholtz Analytic Framework sind Services als Kombination aus Anwendungscontainer, virtuellem Server und On-demand-Speicher herstellbar.
- ▶ Für sehr umfangreiche bzw. datenintensive Auswertungen ist die Nutzung von High Performance Clustern z. B. auf der Basis von Grafikprozessoren (GPU) sinnvoll. Ein solcher Service kann in Verbindung mit Speicherdiensten angeboten und für Forscher bereitgestellt werden, an deren Zentren dann nicht gesondert GPU-Technik aufgebaut werden muss.

Die zweite Gruppe von Services beinhaltet Dienste, die als Plattform für Communities zur Verfügung gestellt werden. Beispiele hierfür sind:

- ▶ Der Themenkomplex Software Services beinhaltet die Ausprägung von Services, Entwicklungsplattformen und Registern im Zusammenhang mit einem professionellen und nachhaltigen Umgang mit wissenschaftlicher Software. Diese Services sind über die Cloud für alle Beteiligten zugänglich und können genutzt werden.
- ▶ Ein anderes Beispiel eines Platform as a Service Layers könnte ein verteiltes Kubernetes Cluster, basierend auf den vorhandenen IaaS Lösungen sein. Dieses Cluster kann containerisierte Community Frameworks dynamisch verteilen, skalieren und verwalten.
- ▶ Ein JupyterHub Service stellt eine Lösung dar, die eine Vielzahl von Single-User Jupyter Notebook Servern z. B. in einem Kubernetes Cluster erzeugen und verwalten kann.
- ▶ Oberhalb von PaaS können auch sogenannte serverless Architekturen angeboten werden. In diesen Modellen werden vorgefertigte containerisierte Funktionen als hochskalierbare Services zur Verfügung gestellt.
- ▶ Im Rahmen des Helmholtz Metadata Center ist die Erstellung und Orchestrierung einer sog. Helmholtz Metadata Fabric vorgesehen, bestehend aus verschiedenen Registern und Diensten zum Umgang mit Metadaten.
- ▶ Im AWI wurde eine Plattform entwickelt für die Registrierung und die Verwaltung von Sensoren. Die Lösung ist aufgrund des bislang verwendeten Beschreibungsstandards SensorML begrenzt auf eine Nutzung in der Erde/Umwelt-Community. Über die Bereitstellung als Cloud-Dienst können die anderen Zentren aus diesem Forschungsbereich das System nutzen, eine Erweiterung auf andere Sensorbeschreibungsstandards wäre prüfbar.
- ▶ Im DKFZ wird daran gearbeitet, in Verbindung mit Object Storage eine Speicherlösung für verteiltes Arbeiten von Radiologen verschiedener Zentren zu entwickeln. Ein solcher Dienst, vorerst begrenzt auf die Radiologen im LifeScience-Bereich, ist ein erster Schritt, um auf dieser Basis auch für andere Communities Lösungen anbieten zu können.

In der dritten Gruppe werden sog. Collaboration Services geführt, also Dienste, die die gemeinsame Arbeit von Wissenschaftlern unterschiedlicher Zentren und Communities unterstützen. Gängige (jedoch teilweise noch ungelöste) Anforderungen sind u. a.:

- ▶ In den verschiedenen Zentren und Communities gibt es eine große Anzahl von sync&share Lösungen, als einer Cloud-Lösung, die das synchronisierte Speichern und Teilen von Daten und Dokumenten erlaubt. Regelmäßige Abfragen im Helmholtz Koordinierungsausschuss Datenverarbeitung (KODA) ergaben allein in der Helmholtz-Gemeinschaft diverse Lösungen, u. a. auf der Basis von OwnCloud, NextCloud oder PowerFolder. Ein zentral von einem Zentrum aufgesetzter Service z. B. auf Basis Next-Cloud entlastet alle Zentren, die einen solchen Service transparent für den Nutzer anbieten wollen, ohne ihn selbst zu betreiben.
- ▶ Die Verbreitung von Google Docs zur browsergestützten, gemeinsamen Bearbeitung von Dokumenten ist der Tatsache geschuldet, dass die gebotene Funktionalität vielfach für die Erstellung von Anträgen, Berichten, Protokollen etc. benötigt wird. Gleichzeitig wird eher selten darauf geachtet, mit welchen Bedingungen die Nutzung verbunden ist. In Verbindung mit NextCloud ist die Nutzung von OnlyOffice mittlerweile funktional so gut geworden, dass es als praxistaugliche Alternative zu Google Docs taugt, bei der auch Datenschutz und IT-Sicherheitskriterien eingehalten werden können. Ein solcher Service für die Helmholtz-Gemeinschaft kann relativ schnell bereitgestellt werden, da verschiedene Zentren schon gute Erfahrungen damit haben (u. a. HZB und DESY).
- ▶ Für die Planung und Durchführung großer Veranstaltungen wie Konferenzen wird in einigen Zentren Indigo genutzt, ein Softwarepaket vom CERN. Insbesondere im Bereich Matter ist diese Software quasi der Standard. Gleichzeitig ist die Wartung und Pflege der Software aufwändig, weshalb auch jetzt schon Zentren eine eigene Installation ablehnen. Als Service würde für diesen Dienst der Aufwand zentral einmal entstehen und die Vorteile der Software wären für viele nutzbar.
- ▶ Im Bereich Software Services wird die Etablierung von Chat-Werkzeugen für die Etablierung dedizierter Kommunikationskanäle thematisiert. Auch für solche Werkzeuge gibt es gute Beispiele in der Helmholtz-Gemeinschaft, so wird die Installation von Rocket Chat im FZJ von den Nutzer sehr gut angenommen.

Bei der Auswahl der Services für ein initiales Portfolio ist besonders bei den Infrastruktur Services die direkte Einbeziehung der Helmholtz Data Federation geplant. Die dort geschaffenen hardwareseitigen Voraussetzungen und die Cloud-Plattform mit seinen Zugriffs- und Servicekomponenten ergänzen sich auf Grundlage des Helmholtz Backbone ideal, so dass im Zusammenspiel der Themen z. B. Speicherdienste der Wissenschaft in der Helmholtz-Gemeinschaft zur Verfügung stehen. Dazu wurde während der Bearbeitung der Proposals mit mehreren Zentren abgestimmt, welche Services für ein initiales Serviceportfolio in Frage kommen und der entsprechende Aufwand in den Aufgabenkomplex Serviceintegration aufgenommen.

Die Entwicklung des Helmholtz Cloud Service Portfolios ist ein kontinuierlicher Prozess. Vor dem Start einer Realisierung bzw. Implementierung weiterer Services in das Portfolio erfolgen vertiefte Evaluierungen auf Basis vorhandener Erfahrungen und aktueller Entwicklungen. Neben der Neuaufnahme von Cloud Services ist auch ein Review-Prozess des bestehenden Portfolios notwendig, der unter anderem Service Level Agreements (SLA) und Operation Level Agreement (OLA) berücksichtigt, die Ergebnisse des Monitoring und Accounting einbezieht und ggf. die Löschung von Services aus dem Portfolio empfiehlt. Das o. a. organisatorische Regelwerk für diese Prozesse wird ebenfalls kontinuierlich überprüft und ggf. angepasst.

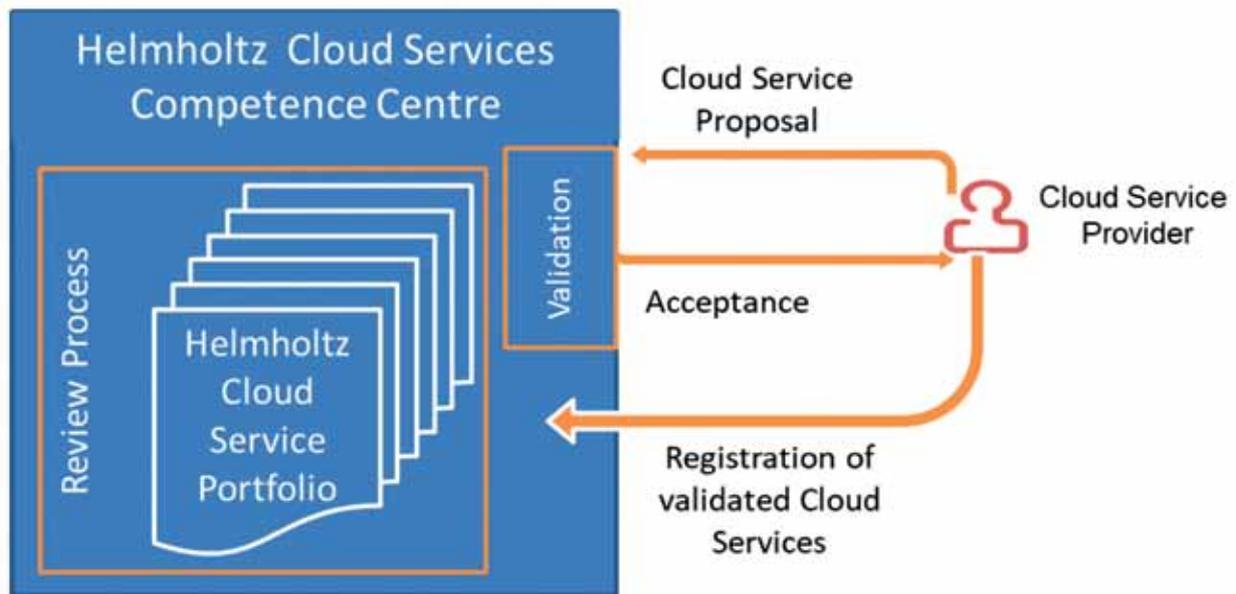


Abbildung 5: Entwicklung des Cloud Service Portfolios

Sowohl für die Erarbeitung des initialen Serviceportfolios als auch in der sich nach dessen Realisierung anschließenden Betriebsphase werden auf Grundlage des Regelwerkes Services bewertet und ggf. für die Integration in das Portfolio vorgeschlagen. Die Entscheidung darüber trifft das Helmoltz ICT Federation Board.

Technische Plattform

Im vorgeschlagenen Konzept erhalten interne und externe Nutzer über den Helmoltz Cloud Access Layer und das Helmoltz Backbone Zugang zu den Cloud Services (siehe auch folgende Abbildung).

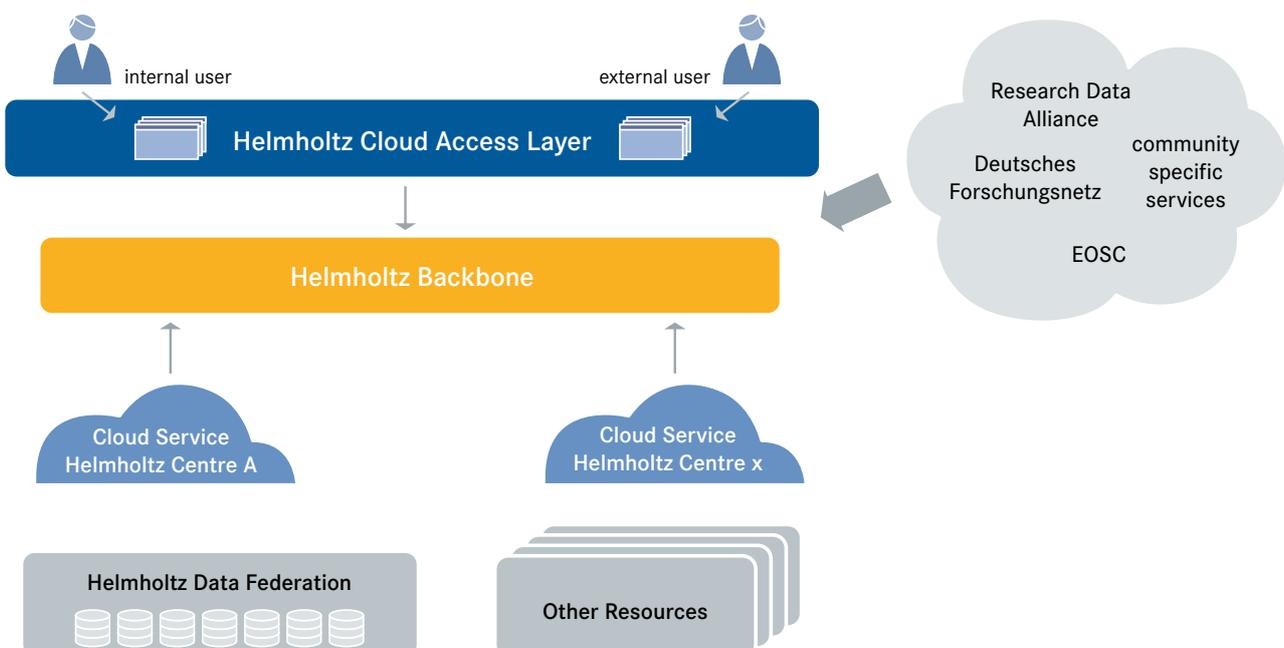


Abbildung 6: Cloud Service Layer

Ein Zugangsportal zu den Helmholtz Cloud Diensten bildet die einheitliche Schnittstelle für alle Nutzer, die solche Dienste in Anspruch nehmen. Es ist damit von besonderer Bedeutung für die Attraktivität der Dienstplattform und für die Stabilität bei der Nutzung. Daraus resultierend ist dem Design, der Funktionalität, der Usability und der Betriebsstabilität besonderes Augenmerk zu widmen.

Aus Designsicht ist zu beachten, dass dieser zentrale Zugang als stand-alone Lösung mit allen Funktionalitäten der Core Services für Zugang, Überwachung der Dienste und ggf. der Abrechnung der Nutzung funktionieren muss, sich aber auch integrierbar zeigt in die Strukturen der Zentren-internen Informationstechnik.

Die Zugangsplattform ist stetig weiterzuentwickeln, um die Erweiterung der Cloud-Dienste abzubilden. Dies erfordert vom Betreiber der Plattform, dass er die IT-Betriebsprozesse optimal gestaltet und so auch sicherstellt, dass die Verfügbarkeit der zentralen Services auf einem hohen Niveau gewährleistet werden kann. Ein entsprechendes Betriebskonzept ist neben der Entwicklung dieser Plattform vom Betreiber zu erstellen und mit dem Cloud Competence Center abzustimmen.

Für die Nutzung der Cloud Services kann die Autorisierung über verschiedene Wege sinnvoll sein:

- ▶ Registrierung eines gesamten Helmholtz-Zentrums (Services werden hauptsächlich transparent vom User direkt vom Zentrum benutzt. z. B. Archiving services),
- ▶ Benutzerbasierte Registrierung als Angehöriger eines Helmholtz-Zentrums,
- ▶ VO (Virtuelle Organisation)-basierend (z. B. Forschungsinfrastruktur),
- ▶ ausschließlich individuell benutzerbasiert (auch für externe Partner möglich).

Für die Authentifizierung sind je nach Integrationstiefe und Servicelevel verschiedene Modelle denkbar wie z. B. Authentifizierung durch Cloud Anbieter oder durch das Cloud Competence Center (CCC). Welches Modell den größten Nutzen für die Forscher und die beteiligten Zentren bietet, kann von Cloud Service zu Cloud Service differieren, unter anderem auch abhängig davon, welche Sicherheitsanforderungen zu erfüllen sind. Die Entwicklung der Genehmigungs-routinen und -allokationen erfolgt in einem vom Cloud Competence Center geleiteten Prozess für jeden Cloud Service bzw. falls sinnvoll für Cloud Service Cluster auf Grundlage des Betriebs- und Steuerungsmodells.

Neben den oben beschriebenen 3 Servicegruppen (IaaS, PaaS und den Kollaborativen Tools) werden noch systemübergreifende Services benötigt, die die Benutzung der Gesamt-Infrastruktur überhaupt erst handhabbar machen.

- ▶ Health und Resource Monitoring: Wie noch später in den Core Services beschrieben wird, ist das Monitoring der Services und der Ressourcen essentiell, um die Funktionsfähigkeit des Gesamtsystems zu jedem Zeitpunkt darzustellen. Monitoring Informationen dienen dazu, bei Störungen zeitnah eingreifen zu können, dem Benutzer einen Überblick über den momentanen Gesundheitszustand des Systems zu geben, was unnötige Supportanfragen verhindern kann und um die Performance und Availability des Systems mit möglichen Service Level Agreements abzugleichen.
- ▶ Professionelles Deployment Management: Da es unerlässlich ist, Software Produkte regelmäßig auf den neuesten Stand zu bringen, nicht zuletzt um sicherheitsrelevante Komponenten zeitnah den jeweiligen, schnell wechselnden Anforderungen anzupassen, ist ein professionelles Produktdeployment von großer Bedeutung. Solch ein System benötigt spezielle technische sowie Prozesskomponenten, wie zum Beispiel eine Testinfrastruktur, einen geordneten Stage Rollout Prozess sowie gut funktionierende Feedback Kanäle zu den Entwicklern.

- ▶ **Application Database:** Es ist einleuchtend, dass Services, die die nötigen Voraussetzungen erfüllen, um im Helmholtz Cloud Service Portfolio gelistet zu werden, auch elektronisch auffindbar und ausreichend gut bezüglich ihrer Funktionalität als auch möglicher Nutzungseinschränkungen beschrieben werden müssen. Weiterhin muss, ohne zusätzliches Suchen, der Benutzer direkt auf die entsprechenden Services umgeleitet werden können.
- ▶ **Grafisches Dashboard:** Eine Erweiterung der Application Database sind personalisierte Dashboards. In der Grundausstattung geben sie dem Benutzer eine Übersicht über die für ihn wichtigen Services, z. B. der ihm zur Verfügung stehende Speicherplatz, der Zustand seiner Serviceanfragen und die Länge von Verarbeitungsschlangen von Computerressourcen. Darüber hinaus muss es dem Benutzer möglich sein, für ihn wichtige Information in sein persönliches Dashboard einzuklinken.

Weitere Core Services sind z. B. Accounting und Monitoring, wobei davon ausgegangen wird, dass bei der Evaluierung eines initialen Cloud-Serviceportfolios noch andere Dienste identifiziert werden, die als Core-Services zu implementieren sind.

Service Integration

Für die verschiedenen Cloud Services wird es sinnvoll sein, die Integrationstiefen und Servicelevel im Access Layer zu differenzieren (siehe auch folgende Grafik). Mögliche Integrationstiefen/ Servicelevel sind dabei z. B.

- ▶ **Betreuer zentraler Katalog** der registrierten Cloud Services (einschließlich der Fähigkeiten und technischen und/oder politischen Beschränkungen und Anforderungen) mit direkter Verlinkung,
- ▶ **Interoperabilität:** wie im EOSC Pilot WP6 Interoperability, weitergeführt von eXtreme DataCloud (Beispiele: Provisionierungssprache z. B. TOSCA, gemeinsame storage protocols wie GridFTP oder WebDAV, gemeinsames Vokabular der Servicequalitäten (QoS)),
- ▶ **Integration:** Nutzung der Interoperabilität, indem Services schon als ‚composed services‘ angeboten werden (Beispiel: compute nodes in Hamburg mit big data storage oder archive in Dresden),
- ▶ **Föderation:** Zusammenschaltung typengleicher Services zu einer Föderation (Beispiele: GRID, EGI fed-cloud, Speicher-Services basierend auf HDF).

Die Entscheidungsgrundlagen für die Realisierung der verschiedenen Integrationstiefen werden zunächst in der Startphase der Plattform entwickelt. Diese Grundlage wird in der operativen Phase der Plattform kontinuierlich weiterentwickelt.

Das CCC entwickelt die Wege für die Bereitstellung des Zugangs der Nutzer zu den Cloud Services wie z. B. Webzugang für die einzelnen Nutzer oder Kopplung mit der Zentren-IT und definierte Schnittstellen. Die Gestaltung des Zugangs und die Lokalisierung der Schnittstelle sind unter anderem eng mit der Frage der Integrationstiefe gekoppelt und muss daher in dem entsprechenden Unterarbeitspaket (siehe oben) mitberücksichtigt werden.

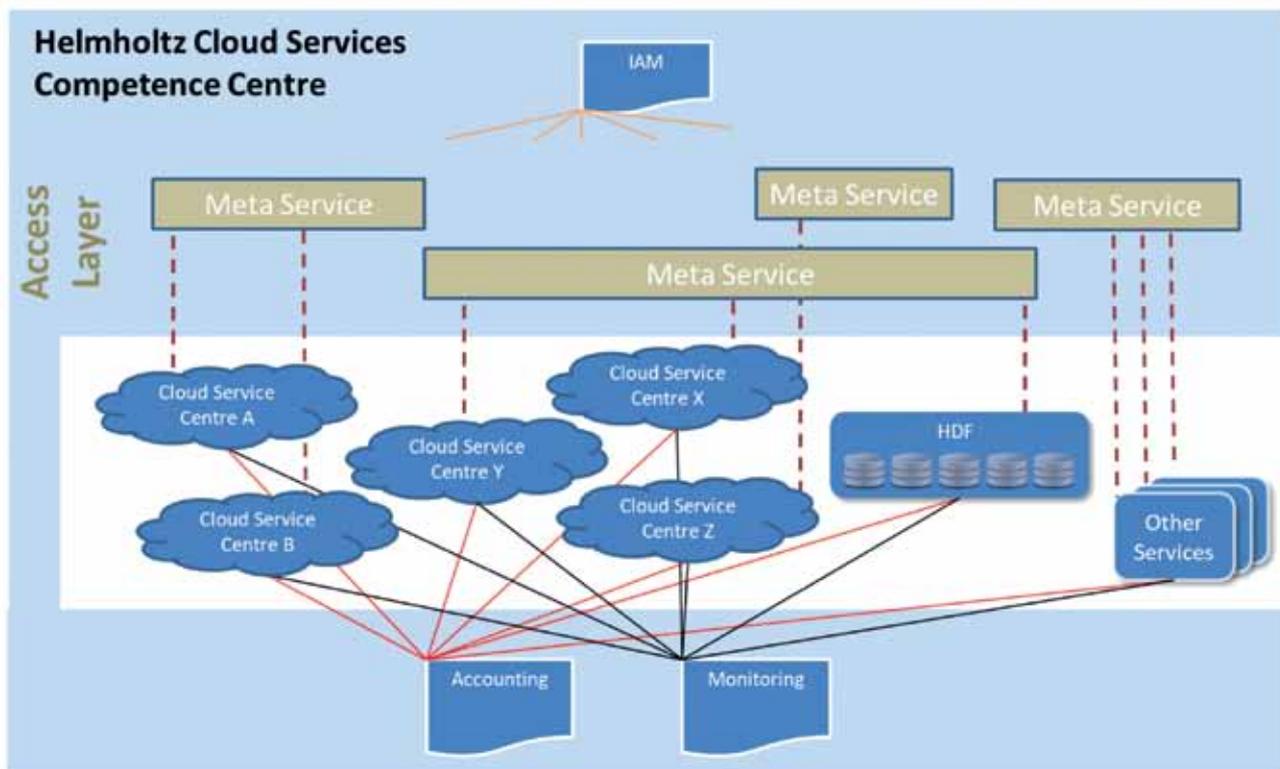


Abbildung 7: Entwicklung des Cloud Service Portfolios

Aus dem vorhergehenden Arbeitspaket Serviceportfolio ist eine Liste von Services erarbeitet, die im Rahmen des Aufbaus der Cloud-Plattform realisiert werden (First Services Integration). Im Rahmen der Arbeitspakete Access-Layer und Core Services sind die technischen Grundlagen soweit erarbeitet, dass die Integration dieser Services realisiert werden kann.

Aufgrund der Heterogenität der fraglichen Infrastruktur-, Plattform-, Scientific-, Software- und Collaboration-Services wird davon ausgegangen, dass die innerhalb der Helmholtz-Gemeinschaft unzweifelhaft vorhandenen Kompetenzen weit verteilt sind. Aus diesem Grund wird dieses Arbeitspaket modular geplant, wobei die Einbeziehung diverser Zentren mit jeweils hervorragendem Know-how bzw. einer existierenden Lösung für einen relevanten Themenbereich über Zielvereinbarungen geplant ist.

Die Phase der Service Integration endet für jeden Service zu unterschiedlichen Zeitpunkten, je nach Umfang der für die Implementierung notwendigen Arbeiten. Die Services werden zu einem definierten Zeitpunkt für die Nutzung freigegeben und gehen damit in die Betriebsphase über.

Betrieb, Support

Zu der Bereitstellung eines Services in der Helmholtz Cloud Plattform gehört eine organisierte Unterstützung der Nutzer und nutzenden Zentren. Dazu gehört neben der Unterstützung bei der Integration von Diensten auch die Bearbeitung und Behebung von Störungen sowie die Beseitigung von Fehlern in der angebotenen Lösung. Der Support für die individuellen Services wird dabei durch den jeweiligen Service Provider sichergestellt.

Von zentraler Stelle aus erfolgt neben der Unterstützung der Nutzer für zentrale Komponenten wie Zugangsportal und dessen Integration in die Zentren-IT auch die Überwachung, Bewertung und stetige Verbesserung der Serviceprozesse für die Cloud Services. Dieser Bestandteil der Erfolgskontrolle wird regelmäßig an das

Helmholtz ICT Federation Board berichtet. Bei bestehenden Defiziten werden nach Abstimmung mit dem Service Provider in diesem Zusammenhang auch Empfehlungen zur Beseitigung unterbreitet.

Governance, Monitoring und Erfolgskontrolle

Die Entwicklung der Cloud Plattform und die Steuerung bzgl. der Integration von Services erfolgen durch das Helmholtz Cloud Kompetenzcluster. In ihm sind die technischen und organisatorischen Kompetenzen und die beteiligten Zentren vertreten, die die Prozesse und Arbeitspakete organisieren und begleiten.

Die Berichtspflichten werden im Kapitel 7 insgesamt für alle drei Plattformkomponenten dargestellt. Der Erfolg der Serviceangebote kann durch folgende Maßnahmen ermittelt werden:

- ▶ Monitoring der Nutzung der bereitgestellten Cloud-Services auf Nutzungs-, Nutzer- oder Zentren-Basis,
- ▶ Monitoring der Support Services für die Cloud-Dienste anhand festgelegter Key Performance Indikatoren (z. B. Anzahl der Störungsmeldungen, durchschnittliche Bearbeitungszeit, ...),
- ▶ Kundenzufriedenheitsanalyse.

5.2 BACKBONE SERVICES

STATUS QUO

Das Deutsche Forschungsnetz (DFN) verbindet Hochschulen und Forschungseinrichtungen miteinander und ist in den europäischen und weltweiten Verbund der Forschungs- und Wissenschaftsnetze integriert. Die folgende Grafik stellt die Anbindung von Helmholtz-Zentren in das XWiN-Glasfasernetz des DFN dar.

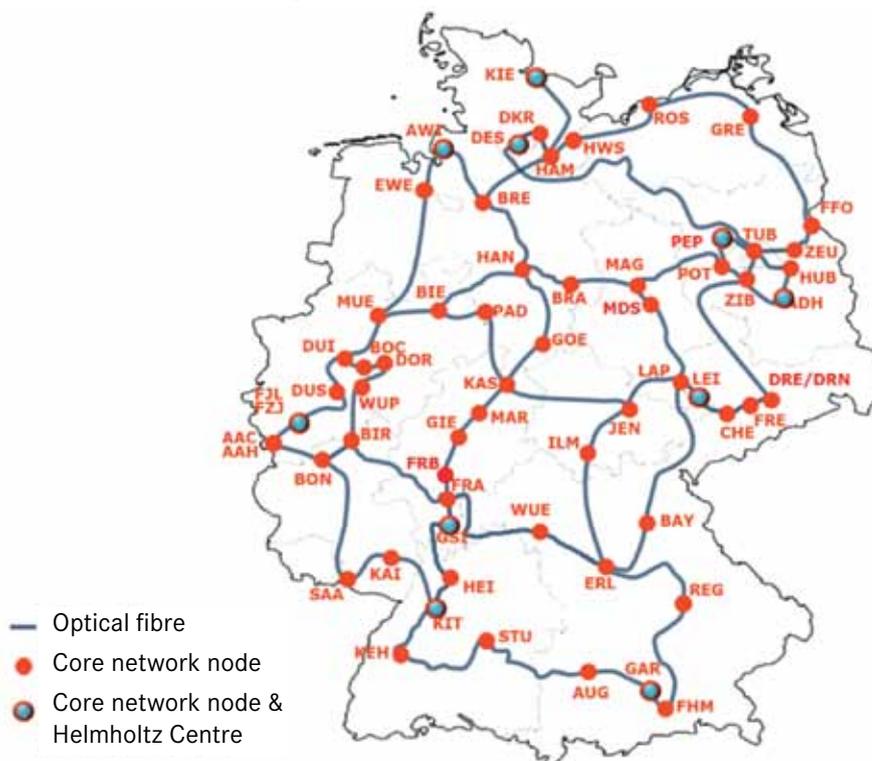


Abbildung 8: Kernnetznoten des DFN X-WiN und Helmholtz-Zentren (Datengrundlage DFN 2018)

Die Zentren der Helmholtz-Gemeinschaft sind an das Deutsche Forschungsnetz (DFN) direkt überwiegend (56 %) mit Geschwindigkeiten zwischen > 1 Gbit/s und < 10 Gbit/s angebunden. 39 % sind mit Geschwindigkeiten von ≥ 10 Gbit/s angebunden⁶. Die folgende Grafik zeigt den Status der XWiN-Anbindung der Helmholtz-Zentren.

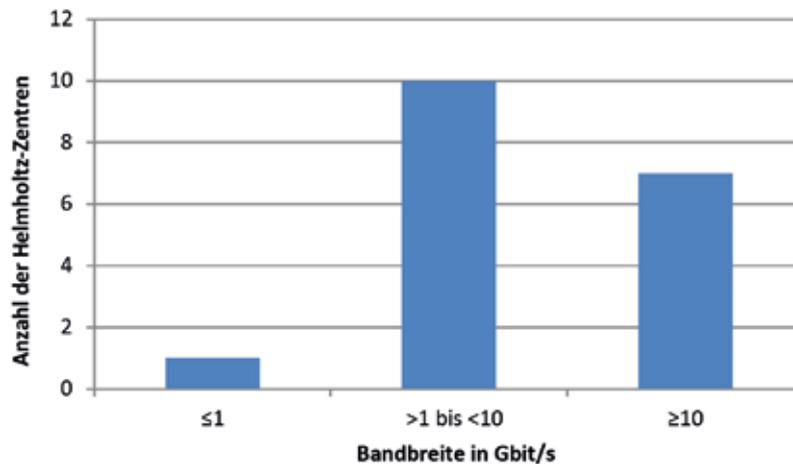


Abbildung 9: Bandbreite des Anschlusses der Helmholtz-Zentren an das XWiN-Netz des DFN (Datengrundlage DFN 2018)

Ein Vergleich mit Forschungsgemeinschaften und den zugehörigen Netzwerken für Forschung und Ausbildung (NREN) in anderen Ländern (siehe auch Analyse zu externem Umfeld im Anhang Kapitel 9.4) zeigt, dass eine hoch performante Vernetzung der großen Labs in einigen NREN schon in den vergangenen Jahren als strategisch für die Forschung identifiziert wurde und hohe Bandbreiten etabliert wurden. Gleichzeitig wurde aber auch deutlich, dass sich die Helmholtz-Zentren in Hinblick auf die Konnektivität in einer vergleichsweise guten Ausgangsposition für die Etablierung eines Helmholtz-Backbones befinden.

Im Rahmen von Workshops wurde mehrfach die Bedeutung ausreichender Bandbreiten für die Forschung der Helmholtz-Gemeinschaft und für die Nutzbarkeit von Cloud Services betont, so unter anderem auch durch externe Gutachter bei der Vorstellung des HIFIS-Konzeptpapiers beim Workshop im März 2018. Dabei wurde auch von den externen Gutachtern betont, dass eine neue Verfügbarkeit von Bandbreite von Forschern für gemeinsame Services und neuen Datenaustausch genutzt wird und somit Bandbreite Datenverkehr erzeugt. Dies deckt sich auch mit den Erwartungen aller NREN, die in dem NREN Compendium 2016 von GEANT befragt wurden (GEANT Compendium 2016).

Durch ein Virtuelles Privates Netzwerk (VPN) kann eine Vertrauensstellung zwischen den beteiligten Zentren effizient genutzt werden. Neben standortübergreifenden Intranets bieten sich VP-Netze auch z. B. für ein inter-institutionelles Datenbackup an, bei dem eine beschränkte Anzahl von Kommunikationspartnern mit einem hohen Datenaufkommen über hochverfügbare Netze verbunden werden. Im XWiN-Netz des DFN sind bereits einige VPN geschaltet (DFN 2018b), an dessen Erfahrungen und Grundlagen angeknüpft werden kann.

⁶ IPP wurde hierbei mit der Anbindung über das MP-Rechenzentrum berücksichtigt, MDC mit der geplanten Bandbreite. Anzumerken ist zudem, dass die Zentren teilweise in weitere Netze wie z. B. Landeswissenschaftsnetze eingebunden sind, beispielsweise in Bremen (Landeswissenschaftsnetz LWN e. V.) oder Berlin (BRAIN – Berlin Research Area Information Network). Teilweise bestehen zudem Datenverbindungen bzw. sind in Planung, die an spezifische Datenbedarfe angepasst sind, wie z. B. die geplante Verbindung zwischen Bremerhaven und Potsdam mit einer geplanten Bandbreite von 10 Gbit/s. Anbindungen von Außenstellen sind hier nicht berücksichtigt.

Ein essentielles Element, um eine Vertrauensstellung zu etablieren, ist eine Struktur zum Management von Identitäten, zur Authentifizierung und zur Autorisierung. Im Deutschen Forschungsnetz (DFN) wurde im Jahr 2007 eine Authentifikations- und Autorisierungs-Infrastruktur (AAI) in den Regelbetrieb übernommen, die einen organisatorischen und technischen Rahmen für den Austausch von Benutzerinformationen zwischen Einrichtungen und Anbietern mit einer notwendigen Vertrauensstellung für die angeschlossenen Wissenschaftseinrichtungen schafft. Zu den Hauptaufgaben der DFN-AAI zählen der technische Betrieb (Lokalisierungsdienst, Testumgebung, Web-Portal mit Informationen), das Erstellen von Richtlinien für die Mitgliedschaft sowie Vertragsgestaltung und -abschluss (DFN 2018). Auch in regionalen Kontexten wurden teilweise AA-Infrastrukturen aufgebaut, die teilweise an die DFN-AAI anknüpfen, wie z. B. in Baden-Württemberg (BWIDM 2018).

Eine europäische Perspektive für ein föderiertes access management für Forschungseinrichtungen wurde z. B. im AARC-Projekt (Authentication and Authorisation for Research and Collaboration) eingenommen. In diesem Projekt wurde mit der AARC Blueprint Architecture (BPA) eine Reihe von Software-Bausteinen vorgeschlagen, mit denen sich föderierte Zugangsmanagementlösungen für internationale Forschungsk Kooperationen implementieren lassen.

Die Entwicklung einer AAI-Lösung muss in jedem Fall die internationale Verflechtung der Forschungsaktivitäten der Helmholtz-Zentren berücksichtigen und eine Kompatibilität zu internationalen Ansätzen schaffen. Hierzu zählt auch die European Open Science Cloud (EOSC) (siehe auch EOSC 2018), EOSC 2018a), die nicht zuletzt aufgrund des open-science-Ansatzes der Helmholtz-Gemeinschaft von wesentlicher Bedeutung ist.

ZIEL

Ziele dieses Arbeitspaketes sind die Bereitstellung einer stabilen Netzwerkinfrastruktur und von gemeinschaftlich nutzbaren Core Services, um dem zunehmenden Bedarf an Vernetzung der Forschung und dem zunehmenden Datenvolumen in der Helmholtz-Gemeinschaft (Stichwort „Datenschatz“) zu entsprechen (siehe hierzu auch die Darstellung in den Science Cases Kapitel 2). So soll der Verbund der einzelnen Helmholtz-Zentren auf Basis einer hohen Bandbreite mit gegenseitigen Vertrauensstellungen und insgesamt erhöhter Sicherheit vernetzt werden. Es umfasst ein Virtual Private Network (VPN), das alle Zentren einschließt und ein gesichertes Qualitätsniveau. Das Netz soll insbesondere die Zentren-übergreifende Nutzung wissenschaftlicher Instrumente und Plattformen ermöglichen sowie zentrale Basis-Services, die für kollaboratives Arbeiten notwendig sind, bereitstellen. Zu diesen Basisdiensten gehört vor allem ein einheitliches Identity & Access Management (IAM). Das physikalische Netz zusammen mit den Basisdiensten bilden die Voraussetzungen und das Fundament für eine Vielzahl höherwertiger Dienste wie z. B. Cloud-Services.

SCHRITTE ZUR ZIELERREICHUNG

Dieses Arbeitspaket strukturiert sich in drei Bereiche:

- ▶ Der Aufbau einer VPN-Struktur zunächst zwischen den Hauptstandorten der Helmholtz-Zentren, möglichst auf der Basis des vom DFN-Verein betriebenen Wissenschaftsnetzes XWiN. Dieses erfordert konkret Vertragsverhandlungen mit dem DFN Verein mit dem Ziel, eine hoch performante VPN-Struktur auf den DFN Plattformen zu schalten. Nota bene, dass es in der Verantwortung der einzelnen Zentren verbleibt, einen eigenen „passenden“ Anschluss am jeweiligen Standort bereitzustellen. Hier treten an

einigen Zentren zusätzliche Kosten auf. Erste Vorgespräche mit dem DFN-Verein haben dazu bereits stattgefunden. Wenn ein solches VPN logisch einmal aufgesetzt wurde, dann ist es operativ in Abstimmung mit dem DFN zu betreiben und zu begleiten. Der Betrieb umfasst dann die Einbindung sich ändernder Netzkonfigurationen, die Kommunikation mit dem DFN Verein sowie die Entwicklung und Anpassung sowie die Auswertung von Monitoring und Accounting Daten, z. B. auf der Basis existierender perfSONAR-Software. Monitoring und Accounting umfassen aber über das Netzwerk hinausgehend alle Ressourcen, die Helmholtz-weit im Rahmen dieses Vorhabens bereitgestellt werden. Diese Informationen dienen dann auch dazu, die Verfügbarkeit der Ressourcen gegenüber den Service Level Agreements (SLA) abzugleichen. Auch die Weiterentwicklung des Konzeptes und die Beratung von Einrichtungen z. B. in der Konfiguration von Firewalls ist Bestandteil des Kompetenzclusters.

- ▶ Neben der Bereitstellung und dem Betrieb einer schnellen Netzwerk-Infrastruktur steht die gemeinschaftliche Nutzung von Core Services im Vordergrund. Hierzu gehört, basierend auf bereits existierenden europäischen Lösungen, ein Helmholtz-weites ausrollen einer Authorisation and Authentication Infrastructure (AAI) sowie eines Identity Managements (IDM), die essentiell für kollaboratives Arbeiten sind. Dieses soll in engem Verbund mit der HDF-Initiative erfolgen und erfordert ein nachhaltiges Betriebskonzept. Weitere wichtige zentrale Dienste liegen z. B. im Nutzersupport für die HIFIS/HDF Services z. B. durch die Bereitstellung und den Betrieb eines Request-Tracker Systems, wie z. B. das am KIT betriebene Global Grid User Support (GGUS).
- ▶ Von besonderer Bedeutung sind auch Daten Transfer Services, die als Core Service angeboten und betrieben werden sollen. Auch die Bereitstellung von gegenseitigen Backup Möglichkeiten zwischen Zentren führt zu erheblicher Betriebssicherheit und hilft, Helmholtz-weit Kosten zu sparen.

Alle Dienste sollen über alle Arbeitspakete hinweg in Katalogen erfasst und dem Nutzer bekannt gemacht werden. Sie unterliegen alle einem überwiegend professionellen Deployment.

VPN

In der Startphase wird der Aufbau eines logischen Datennetzwerks (voll vermaschtes Virtual Private Network (VPN)) mit der anfänglichen Zielmarke von mindestens 10 Gbit/s durch das „Backbone Services Competence Centre“ (BCC) initiiert, organisiert und begleitet. Ein zügiger weiterer Ausbau auf 40 Gbit/s und dann 100 Gbit/s wird zunächst zwischen den Hauptstandorten der Helmholtz-Zentren avisiert. Das LHCONE⁷ kann dabei als Blaupause für das geplante Helmholtz-VPN angesehen werden. Abbildung 9 zeigt, dass bei 39 % der Zentren ein Anschluss mit der avisierten Startbandbreite von mindestens 10 Gbit/s bereits besteht.

Der Aufbau des Netzes erfolgt vorzugsweise auf Basis einer DFN Partnerschaft. Grundlagen des Helmholtz-Backbones bilden demnach die Zentrenanschlüsse in das XWiN. Es ist vorgesehen, dass der entsprechende leistungsfähige Anschluss eines Helmholtz-Zentrums an das Helmholtz-Backbone von jedem Helmholtz-Zentrum eigenständig erbracht wird, die Zentrums-übergeordneten Elemente jedoch über den Inkubator realisiert werden.

Es ist prioritär eine Vernetzung der Haupt-Zentrumsstandorte geplant, eine Einbeziehung weiterer Standorte bzw. Außenstellen kann sich ergeben. Das Helmholtz-Netzwerk soll so aufgebaut werden, dass es offen für externe Nutzer (insbesondere Kooperationspartner der Forscher) ist. Hierfür sind dann entsprechende Sicherheitsmaßnahmen zu entwickeln und zu implementieren. Das im Rahmen der Antragsstellung

7 siehe z. B. <http://lhcone.web.cern.ch/>

erarbeitete Konzept muss in Kooperation mit den beteiligten Partnern kontinuierlich weiterentwickelt und detailliert werden.

Die vertraglichen Grundlagen werden in Zusammenarbeit mit den beteiligten Partnern durch das BCC entwickelt. Der technische Aufbau erfolgt für die Zentrums-übergeordneten Aspekte durch den gewählten Netzwerk-Anbieter (vorzugsweise DFN) und für die Zentrums-bezogenen Teile durch die jeweiligen Zentren, für die ggf. Unterstützung durch das BCC geleistet wird.

Als Teil des übergeordneten Monitorings und Accountings wird auch das Helmholtz Backbone durch entsprechende Tools wie z. B. perfSONAR regelmäßig untersucht, z. B. um Verkehrsbeziehungen zwischen den einzelnen Zentren bei z. B. einem gemeinsamen Betrieb von Instrumenten zu analysieren und basierend darauf Planungen vorzunehmen.

Wesentlich im Rahmen der Entwicklung eines Helmholtz Backbones ist ebenfalls die Sicherstellung einer definierten Quality of Service. Dies umfasst z. B. die Realisierung garantierter Bandbreite zwischen den Zentren, um insbesondere wissenschaftliche Experimente remote betreiben zu können. Hierzu wird das BCC die Entwicklung von Prozessen und Mindestanforderungen und das Monitoring der Realisierung organisieren. Ein wichtiger Teil der Definition geeigneter Prozesse wird auch die Prüfung der Möglichkeit eines einheitlichen Namensraums und einheitlicher Attribute (REFEDS Standard) sein.

VERTRAUENSSTELLUNG, IDENTITÄTSMANAGEMENT

Durch den Aufbau von Vertrauensstellungen, die leichteren Zugang zu Ressourcen untereinander ermöglichen und dem gleichzeitigen Ziel von erhöhter IT-Sicherheit, wird ein geringerer betrieblicher Aufwand für Schutzmechanismen innerhalb der Helmholtz-Gemeinschaft erreicht. Hierdurch können wesentliche Synergien mobilisiert werden. Ein zentrales Element stellt ein Helmholtz-weites Identity & Access Managements (IAM) dar. Dabei werden die Identitäten von Anwendern aus den Zentren so gefördert, dass sie für die Services einheitlich verwendet werden können und der Anwender mit dem gleichen Benutzernamen und Passwort auf alle Services (z. B. Datenzugriff, Datenanalyse, Computerressourcen, Kollaborationstools, etc.) zugreifen kann, sofern sie für ihn freigeschaltet sind.

Bei der Entwicklung können und sollten die Erfahrungen in anderen Wissenschaftsnetzwerken aufgenommen und berücksichtigt werden. Hier sind beispielsweise die Entwicklungen in dem regionalen Netzwerk von Baden-Württemberg zu nennen.

Über eine Helmholtz-weite harmonisierte Authentication and Authorisation Infrastructure (AAI) erfolgt eine Authentifizierung und Autorisierung von Benutzern bzw. Benutzergruppen für bestimmte Services in der Helmholtz-Gemeinschaft. In der Initiative OFFDI (OFFDI 2018) wurden verschiedene Optionen für einen solchen Dienst evaluiert. Deren Ergebnisse werden bei der Entwicklung und Implementierung aufgenommen.

Die genannten Core Services (AAI, IAM) sollen nicht nur die Basis für Helmholtz-weite Services sein, sondern auch Elemente der Mitarbeit der Helmholtz-Zentren in gemeinschaftsübergreifenden Projekten, in die gerade die Helmholtz-Gemeinschaft sich mit den großen Infrastrukturen als starker Anbieter einbringen muss. Bei der Entwicklung und Implementierung ist es daher wesentlich, existierende Lösungen z. B. von den EU geförderten Projekten AARC2 (AARC 2018), EOSC-HUB (EOSC-HUB 2018) und weiteren zu berücksichtigen, um sich zukünftig auch in europäische und internationale Infrastrukturen integrieren zu können. Eine AAI, die in die DFN AAI integriert ist, kann z. B. als Option für eine optimierte Einbindung externer wissenschaftlicher Partner geprüft werden, sofern sie auch auf europäischer Ebene einsetzbar ist. Wie bei AAI können auch bei IAM

Erfahrungen von Helmholtz-Zentren in den Aufbau der HIFIS-Plattform eingebracht werden, wie z. B. aus den Arbeiten zum Aufbau des Landesnetzes Baden-Württemberg (bwIDM siehe BaWü 2018) oder die Beteiligung an den Arbeiten im Rahmen von INDIGO (Indigo 2018).

Gerade für die Helmholtz-Gemeinschaft wird es wichtig sein, neben den web-basierten Diensten (Einbindung vielfältig realisiert z. B. SAML und Shibboleth) auch die Einbindung von wissenschaftlichen nicht-web-basierten Diensten (z. B. HPC-Ressourcen) zu bearbeiten.

Der Aufbau von Vertrauensstellungen erfordert Policies, die von den teilnehmenden Zentren gemeinsam verabschiedet und weiterentwickelt werden. Dabei wird voraussichtlich nicht jedes Zentrum alle einzelnen Maßnahmen mittragen können. So sind Schutzbedarfe im Gesundheitsbereich anders als z. B. in MATTER. Hier ist ein abgestuftes Konzept vorgesehen und die Hoheit über die Implementierung der Policies verbleibt in den Zentren.

Die Zentren-übergeordnete Entwicklung und Implementierung wird vom BCC in Zusammenarbeit mit den beteiligten Partnern realisiert. Das BCC berät die Zentren bei den Zentren-bezogenen Arbeiten (z. B. IdP-Implementierung). Eine enge und intensive Einbindung der HDF Aktivitäten ist vorgesehen und insbesondere über die sowohl an HDF als auch an HIFIS beteiligten Personen gesichert. Ziel ist es, auch hier die bereits geleisteten Arbeiten zu nutzen und Doppelentwicklungen zu vermeiden.

WEITERE SERVICES

Als weiteren core service wird das BCC die Entwicklung von Daten Transfer Services für die Gesamtheit der Helmholtz-Zentren organisieren. Dabei wird auch hier soweit möglich auf etablierte bzw. bewährte Tools aufgebaut. Zu erwähnen ist hier z. B. der File Transfer Service (FTS), der am CERN entwickelt wurde und als GRID Data Transfer Service am CERN den Großteil der Large Hadron Collider (LHC)-Daten über das World LHC Computing Grid verteilt. FTS unterstützt verschiedene Protokolle und verschiedene Storage Systeme wie z. B. das deutsche dCache.

Die Bereitstellung von gegenseitigen Backup Möglichkeiten zwischen Zentren bildet einen weiteren core service. Bestehende und in der Entwicklung befindliche Backup-Lösungen wie z. B. beim italienischen Istituto Nazionale di Fisica Nucleare (National Institute for Nuclear Physics INFN) werden mit Blick auf die Übertragbarkeit auf die Gemeinschaft der Helmholtz-Zentren evaluiert (u. a. Kompatibilität, Anforderungen an die Infrastruktur, Zuwachs an Betriebssicherheit, Kosteneffekte) und ein Helmholtz-weites inter-institutionelles Backup konzeptioniert und pilotiert. Für die Planung einer Nutzung in der Fläche werden unter anderem bestehende Infrastrukturen der Zentren (z. B. auch Bandbreite) berücksichtigt.

Das BCC organisiert weiterhin die Entwicklung operativer Services, wie z. B. das Helpdesk, Support und Service Management, die im föderierten Ansatz den Nutzern zur Verfügung gestellt werden.

GOVERNANCE, MONITORING UND ERFOLGSKONTROLLE

Die Entwicklung der Backbone Services für die Gemeinschaft der Helmholtz-Zentren wird durch das Helmholtz Backbone Kompetenzcluster realisiert. In ihm sind die technischen und organisatorischen Kompetenzen und die beteiligten Zentren vertreten, die die Prozesse und Projekte technisch und vertraglich fördern, organisieren, realisieren und begleiten (zu weiteren Aspekten der Governance wie z. B. dem Reporting siehe Kapitel 7).

Monitoring und Accounting umfassen nicht nur die Netzwerk- und Core Services. So muss – und dies gilt für alle Teilplattformen – z. B. auch die Verfügbarkeit von Ressourcen überwacht werden und durch Portale, z. B. Dashboards, für den Nutzer sichtbar sein. Ähnliches gilt auch für Services, die den Nutzern zur Verfügung stehen. Diese müssen in Katalogen erfasst und z. B. über Web-Portale veröffentlicht werden.

In der Startphase dieses Arbeitspaketes wird durch das Helmholtz Backbone Competence Centre (BCC) in Zusammenarbeit mit den beteiligten Institutionen die Entwicklung eines Modells für das Accounting und Monitoring (inkl. Erfolgskriterien und -indikatoren) organisiert. Hierbei werden Erfahrungen aus Projekten mit ähnlichen Zielsetzungen ausgewertet und ggf. übertragen (z. B. Modell Hessen).

Im Zusammenhang mit der Sicherstellung einer definierten Quality of Service (s. o.), wie z. B. der Realisierung garantierter Bandbreite zwischen den Zentren, wird das BCC die Entwicklung von Prozessen und Mindestanforderungen sowie das Monitoring der Realisierung organisieren.

5.3 SOFTWARE SERVICES

STATUS QUO

Software ist eine wesentliche Grundlage für den modernen Forschungsprozess. Sie legt die Basis für den langfristigen, stabilen Betrieb der Forschungsinfrastrukturen, stellt gleichzeitig einen wesentlichen Innovationstreiber dar und trägt wesentlich zum Erkenntnisgewinn sowie zur Wertschöpfung bei. Schließlich ist Software selbst ein wichtiges Ergebnis des Forschungsprozesses. Daher ist es nicht überraschend, dass an den Helmholtz-Zentren Software in großer Vielfalt genutzt und entwickelt wird. Das Spektrum reicht von wenigen Codezeilen für spezielle Aufgaben, über Skripte zur Datenaufbereitung und kleinen Software-Werkzeugen, bis hin zu umfangreichen, komplexen Programmpaketen. Zudem drückt sich diese Vielfalt in der Vielzahl verwendeter Programmiersprachen, Technologien und Umgebungen aus.

Die Vielfalt und insbesondere der Umgang mit wissenschaftlicher Software variieren nach Forschungsbereich, Community oder Zentrum stark und richten sich momentan vorrangig nach der Bedeutung für die eigene wissenschaftliche Wertschöpfung und nach den Vermarktungsaussichten. Die Betrachtung von Software als ein eigenständiges Forschungsprodukt und der nachhaltige Umgang damit findet selten statt. Dadurch wird die Nachvollziehbarkeit und Reproduzierbarkeit von Forschungsergebnissen gefährdet (Howinson 2016).

Die Problematik der Nachvollziehbarkeit und Reproduzierbarkeit von Forschungsergebnissen ist der Forschungsgemeinschaft bewusst. Daher hat sich eine Reihe von internationalen und nationalen Initiativen in diesem Kontext herausgebildet, die zunehmend an Bedeutung gewinnen. Im Folgenden geben wir dazu einen kurzen Überblick, im Anhang sind weitergehende Informationen zu diesen Initiativen zu finden:

- ▶ Das Software Sustainability Institute (SSI)⁸ wurde 2010 in Großbritannien gegründet. Es verfolgt das Ziel, die Nachhaltigkeit und die Qualität von Forschungssoftware in der britischen Forschungsgemeinschaft zu fördern und zu verbessern. Aktuell befindet sich ein amerikanisches Research Software Sustainability Institute⁹ nach diesem Vorbild in der Planungsphase. Zudem existiert mit dem eScience Center¹⁰ ein vergleichbares Institut in den Niederlanden.

8 vgl. <https://www.software.ac.uk/about>

9 vgl. <http://urssi.us/>

10 vgl. <https://www.research-software.nl/about>

- ▶ *The Carpentries* ist eine internationale Community mit dem Ziel, Forscherinnen und Forschern essentielle Kenntnisse in den Bereichen Software-Entwicklung und Datenwissenschaften zu vermitteln. Ein wesentlicher Ursprung von *the Carpentries* ist *Software Carpentry*, welche seit 2012 bereits weltweit hunderte Veranstaltungen für mehr als 34.000 Teilnehmerinnen und Teilnehmer durchgeführt hat¹¹.
- ▶ Die *Research Software Engineers Association* ist das Forum der britischen Research Software Engineers (RSE). Die RSE Association organisiert seit 2016 eine jährliche Konferenz, die zunehmend an Bedeutung gewinnt. Es haben sich eine Reihe nationaler RSE-Vereinigungen gegründet¹², aktuell befindet sich der deutsche Ableger (de-RSE¹³) in der Gründung.
- ▶ Im Rahmen der Open Science Initiative der Helmholtz-Gemeinschaft wurde Ende 2016 die Task Group *Zugang zu und Nachnutzung von wissenschaftlicher Software* gegründet. Die Arbeitsgruppe hat bisher ein Positionspapier¹⁴ veröffentlicht sowie konkrete Empfehlungen¹⁵ an die Helmholtz-Zentren zur Schaffung von Leit- und Richtlinien zum Umgang mit wissenschaftlicher Software formuliert und erarbeitet aktuell eine Muster-Richtlinie.
- ▶ Auch über die Grenzen der Helmholtz-Gemeinschaft hinweg gewinnt das Thema Forschungssoftware innerhalb der Deutschen Forschungsgemeinschaft an Bedeutung. In der Schwerpunktinitiative Digitale Information wurde auf Betreiben der Helmholtz-Gemeinschaft für die neue Periode 2018–2022 ein Handlungsfeld „Digitale Werkzeuge – Software und Dienste“ gegründet.

Die aufgeführten Initiativen lassen erkennen, dass die internationale Forschungsgemeinschaft und auch die Helmholtz-Gemeinschaft sich der Bedeutung von Software für die Nachvollziehbarkeit und Reproduzierbarkeit von Forschungsergebnissen bewusst ist. Konkrete praktische Angebote und Plattformen, um die Forscherinnen und Forscher bei deren Umsetzung zu befähigen, fehlen allerdings in der deutschen Forschungslandschaft weitgehend bzw. sind in den Zentren sehr unterschiedlich ausgeprägt.

ZIEL

Ziel der Komponente „Software Services“ ist es, die Helmholtz-Zentren beim nachhaltigen Umgang mit Forschungssoftware durch praktische Angebote und Services zu unterstützen. Konkret sollen die folgenden Aspekte durch das Arbeitspaket nachhaltig adressiert werden:

- ▶ Forscherinnen und Forscher sollen ausgebildet und unterstützt werden, um sie zur Entwicklung und Publikation von qualitativ angemessener, nachhaltiger Software zu befähigen.
- ▶ Den Forscherinnen und Forschern soll eine technologische Infrastruktur für die Planung, die Entwicklung, die Qualitätssicherung, die Publikation, und die Auffindbarkeit von Forschungssoftware dauerhaft bereitgestellt werden.
- ▶ Die Nachnutzung vorhandener Software durch Forscherinnen und Forscher soll verbessert werden, indem Werkzeuge zum Finden von geeigneten Software-Lösungen bereitgestellt werden, Ansprechpartner bzw. Know-how-Träger vermittelt werden sowie Unterstützung beim Aufbau von Communities für konkrete Forschungssoftware geleistet wird.

11 vgl. <https://software-carpentry.org/about/>

12 vgl. <https://researchsoftware.org/>

13 vgl. <http://www.de-rse.org>

14 vgl. <https://os.helmholtz.de/index.php?id=2766>

15 vgl. <https://os.helmholtz.de/index.php?id=3197>

Die Arbeitspakete konzentrieren sich dabei vor allem auf Software Services, die nicht ohne weiteres an einem Zentrum allein erbracht werden können und ein möglichst hohes Synergiepotenzial zwischen den Zentren aufweisen. Neben den über die Competence Centre bereitgestellten Services und Plattformen werden über Arbeitspakete schrittweise weitere Gruppen in den Zentren einbezogen und dadurch gezielt eine breite Expertise und verteilte Supportstrukturen entwickelt.

SCHRITTE ZUR ZIELERREICHUNG

Im Forschungsprozess spielt die Nutzung und die Entwicklung von Software an verschiedenen Stellen eine zunehmende Rolle. Dazu einige Beispiele:

- ▶ Zu Beginn eines Forschungsprojektes kann es notwendig sein, mit verschiedenen Software-Tools und eigenen Skripten zu experimentieren. Hier sind gute Kenntnisse der üblichen Werkzeuge der jeweiligen Domäne relevant. Zudem muss eine ausreichend flexible IT-Infrastruktur existieren, um eine einfache Nutzung von Programmpaketen zu ermöglichen. Nachdem der Ansatz, die Vorgehensweise und der Technologie Stack sich grundlegend stabilisiert haben, sind zumindest die Verwendung eines Versionskontrollsystems und die sichere Ablage in einer zentralen Softwareentwicklungsplattform notwendig. Erst dadurch wird die gemeinsame Weiterentwicklung und Zusammenarbeit ermöglicht.
- ▶ Bei der Publikation der Forschungsergebnisse wird zunehmend gefordert, dass verwendete Daten und Software ebenfalls publiziert bzw. referenziert werden. Dadurch sind selbst im Fall von kleineren Skripten verschiedene Aspekte zu klären: Auf welcher Plattform soll die Software veröffentlicht werden? Welche Lizenz ist strategisch sinnvoll bzw. rechtlich erforderlich? Wie muss die Software aufbereitet und dokumentiert sein, so dass sie richtig referenziert wird bzw. die damit erzielten Ergebnisse nachvollzogen werden können?
- ▶ Sobald eine entwickelte Software eine bestimmte Relevanz und Nutzen in einer Forschungsgruppe zeigt, kann sich diese zu einem wichtigen Forschungs-Framework für die Community entwickeln. Dies geht i.d.R. mit einer langfristigen Weiterentwicklung und Pflege der Software einher. Um eine nachhaltige und effiziente Weiterentwicklung und Bereitstellung der Software zu gewährleisten, ist der Einsatz moderner Entwicklungspraktiken erforderlich. Diese wirken effektiv der durch Erweiterungen verursachten ständigen Zunahme der Komplexität und dem Wissensverlust entgegen, der beispielsweise durch Mitarbeiterfluktuation entsteht.

Für die Helmholtz-Zentren ergibt sich hieraus die Herausforderung, praktikable Regelungen zu etablieren und erforderliche Services in den Bereichen Infrastruktur, Ausbildung und Beratung effizient bereitzustellen. Dafür existiert ein hohes Synergiepotenzial, welches dieses Arbeitspaket durch die Bereitstellung verschiedener Software Services adressiert. Dabei wird die Strategie verfolgt, sich möglichst auf existierende Standards, Initiativen, Services und Infrastrukturen zu stützen. Zudem liegt der Fokus auf der effektiven Nutzbarmachung der Services für eine möglichst hohe Anzahl von Helmholtz-Zentren. Dies ist zugleich eine der zentralen Herausforderungen dieses Arbeitspakets. Daher wird die Schaffung einer nachhaltig wirksamen Organisationsstruktur angestrebt, um die Zentren bei der Ausgestaltung der Services aktiv einzubeziehen und somit eine hohe Nutzungsakzeptanz zu erreichen.

Neben Services, die Entwickler (Wissenschaftler, Doktoranden, Ingenieure und Techniker) von Software in die Lage versetzen, selbst nachhaltige Lösungen zu schaffen, sind auch technologische Services vorgesehen, die geeignete Infrastrukturen für die Planung, den Test und die Publikation bereitstellen. Dafür werden vier

Gruppen von Services entwickelt, die Arbeitspakete der Komponente „Software Services“ darstellen und im Folgenden beschrieben werden.



Abbildung 10: Übersicht über die geplanten SW-Services

Support Services

Die Support Services stellen konkrete Anlaufstellen für Forscherinnen und Forscher im Fall von Fragen und Problemen dar. Zudem ist es vorgesehen, einige Support Services teilweise in Form von vor-Ort-Consulting zu etablieren. Insbesondere letztere Variante verfolgt das Ziel, mittelfristig das Know-how direkt in den unterstützten Forschungsgruppen zu etablieren. Im Folgenden werden die aktuell vorgesehenen Support Services kurz beschrieben:

- ▶ **Anforderungsanalyse und Usability:** Dieser Support Service hilft Forschungsgruppen bei der Analyse der Anforderungen und der relevanten Nutzergruppen, der Erstellung des dazugehörigen Nutzungskonzepts sowie bei dessen Umsetzung.
- ▶ **Architektur- und Software-Analyse:** Dieser Support Service unterstützt Forschungsgruppen bei der Etablierung einer angemessenen Architektur und deren Dokumentation. Zudem werden Code-Analysen angeboten, um Verbesserungspotenziale zu identifizieren und Rückschlüsse auf die Wartbarkeit der Software zu ziehen.
- ▶ **Entwicklungsprozesse und Automatisierung:** Dieser Support Service unterstützt Forschungsgruppen bei der Analyse und der schrittweisen, strategischen Verbesserung ihres Entwicklungsprozesses und dessen Automatisierung (z. B. Testpraxis, Buildprozess, Deployment).
- ▶ **Technologischer Support:** Dieser Support Service unterstützt Nutzerinnen und Nutzer der etablierten technologischen Services. Zudem werden dadurch die Zentren bei der Einführung und dem Betrieb von lokal in den Zentren installierten technologischen Services unterstützt.

- ▶ **Open Source und Inner Source:** Dieser Support Service unterstützt Forscherinnen und Forscher beim Aufbau von Nutzer- und Entwickler-Communities für konkrete Forschungssoftware innerhalb und außerhalb des Zentrums.
- ▶ **Lizenzen und Rechtliche Aspekte:** Dieser Support Service konzentriert sich auf die rechtliche Beratung im Rahmen der Software-Entwicklung. Dazu zählen beispielsweise die Bereiche Urheber-, Lizenz-, Patent- und Exportkontrollrecht sowie Datenschutzgrundverordnung.

Im Fall der Support Services wird prinzipiell angestrebt, dass jeder Service durch ein Kompetenzcluster mit entsprechender Expertise für alle Zentren erbracht wird. Alternativ ist auch in Betracht zu ziehen, bestimmte Support Services ganz oder teilweise durch externe Anbieter erbringen zu lassen.

Technologische Services

In diesem Teilbereich liegt der Fokus auf der nachhaltigen Bereitstellung einer gut integrierten und leicht nutzbaren Infrastruktur, welche die Planung, die Entwicklung, die Qualitätssicherung, die Publikation und die Auffindbarkeit von Forschungssoftware unterstützt. Solche Plattformen sind essentiell, um den Forscherinnen und Forschern eine möglichst leichte Umsetzung der Empfehlungen zu ermöglichen. Wie die vergangene Entwicklung bei Publikationen und bei Forschungsdaten zeigt, ist es essentiell, die eigenen Ergebnisse der Forschung und Entwicklung selbst in der Hand zu behalten und nicht unbedacht privaten Firmen zu überlassen. Die Übernahme von GitHub durch Microsoft im Juni 2018 beweist, wie schnell eine innovative und vermeintlich freie Lösung am Ende bei gewinnorientierten Konzernen landen kann und man damit u. U. sein geistiges Eigentum abgibt.

Beispiele für die vorgesehenen technologischen Services sind:

- ▶ Web-basierte Plattformen zur Unterstützung der kollaborativen Softwareentwicklung, wie z. B. GitHub oder GitLab, stellen eine Vielzahl guter Werkzeuge zur Planung, Entwicklung und der Qualitätssicherung von Software bereit. Zu diesen Werkzeugen gehören i.d.R. Versionskontrollsysteme wie Git, Werkzeuge zur Aufgabenverwaltung und Planung von Software-Releases, Werkzeuge zur Dokumentation wie Gitlab-Pages oder Wikis, Werkzeuge zur Unterstützung für die kontinuierliche Integration, Werkzeuge zur Durchführung von Code Reviews sowie Mechanismen zur Vereinfachung der räumlich getrennten Zusammenarbeit (u. a. Continuous Integration). Diese Werkzeuge werden in einer gut integrierten Form den Nutzern bereitgestellt. Zudem erlauben die Plattformen auch die Integration externer Services, wodurch sie flexibel an den konkreten Entwicklungsablauf angepasst werden können. Eine solche Web-basierte Plattform stellt einen der technologischen Kern-Services in diesem Arbeitspaket dar.
- ▶ Persistente Digital Object Identifier (DOI) sind für die Kennzeichnung verwendeter Software innerhalb einer Publikation unerlässlich. Daher ist es notwendig, dass die bereitgestellte Web-basierte Plattform zur Software-Entwicklung eine direkte Integration mit DOI-Vergabe-Services wie beispielsweise Zenodo für einzelne Software-Releases bereitstellt.
- ▶ Persistente Chat-Werkzeuge wie Slack oder Mattermost erlauben es, einen dedizierten Kommunikationskanal für die Entwicklung einer Software zu etablieren. In diesem Kanal laufen Diskussionen zwischen den Entwicklerinnen und Entwicklern sowie aktuelle Entwicklungsinformationen (z. B. Code-Änderungen, fehlgeschlagene Software-Builds) zusammen. Ein solcher Kanal erleichtert insbesondere die verteilte Zusammenarbeit an einer Software und kann auch von den End-Anwendern genutzt werden.

- ▶ Die kontinuierliche Integration und das Deployment von Software erfordert die Bereitstellung der Ablaufumgebung, welche die notwendigen Abhängigkeiten zum Erstellen und Testen der Software sowie für deren Betrieb enthält. Zur automatischen Bereitstellung werden beispielsweise Container-Technologien oder Werkzeuge zur automatischen Provisionierung eingesetzt. Die resultierende Umgebung wird typischerweise auf Basis einer Cloud-Infrastruktur wie GitLab CI-Runner oder die kostenpflichtigen Amazon Web Services erstellt. Zur Bereitstellung einer solchen Cloud-Infrastruktur bietet sich eine Kooperation mit der Komponente Cloud Services an.
- ▶ Jupyter Notebook / Jupyterlab kombinieren das interaktive Experimentieren mit Code und das Einbetten von Dokumentation. Sie sind gut geeignet, um Neulinge in die Programmierung einzuführen, und werden häufig zur Exploration von Datensätzen eingesetzt. Zur effizienten Nutzung von Jupyter Notebooks, insbesondere zur Verarbeitung großer Datenmengen, ist ein Cloud-basiertes Angebot essentiell. Dieser Service stellt einen Mehrwert insbesondere für die Inkubator Plattform HMC (Metadatenmanagement) dar und erlaubt eine Kooperation mit der Komponente Cloud Services. Darüber hinaus bietet die Jupyter Technologie die Erstellung „lebendiger“ interaktiver Publikationen.
- ▶ Zum Überblick über und dem Auffinden von Forschungssoftware der Helmholtz-Zentren soll ein zentraler Software-Katalog bereitgestellt werden. Der Katalog soll automatisiert anhand der Informationen und Metadaten der Software-Repositoryen der Zentren aufgebaut werden. Die Software-Repositoryen müssen dazu mit einem Metadatensatz ausgezeichnet werden, der insbesondere das Thema Software-Publikation adressiert und die Nachnutzung vereinfacht. In diesem Zusammenhang bietet sich eine Kooperation mit dem Inkubator Plattformantrag HMC an. Zudem bietet ein solcher Katalog die Möglichkeit, die Umsetzung von Empfehlungen zur nachhaltigen Software-Entwicklung beispielsweise über Badges¹⁶ zu kennzeichnen. Eine solche Kennzeichnung stellt eine Form der Anerkennung dar und liefert einen zusätzlichen Anreiz, Empfehlungen umzusetzen.

Ergänzungen der Services sowie Details der Ausgestaltung und der Bereitstellung sollen durch das HIFIS ICT Federation Board mit möglichst enger Einbindung aller Zentren im Verlauf festgelegt werden. Im Fall der Bereitstellung soll jeweils geprüft werden, ob ein technologischer Service durch die Nutzung eines existierenden Angebots, als selbst-betriebener Cloud-Service innerhalb der HIFIS-Plattform oder in Form lokaler Installationen in den Zentren zweckmäßig und wirtschaftlich etabliert werden kann. Zudem ist im Entscheidungsprozess darauf zu achten, eine zu hohe Abhängigkeit von einzelnen Anbietern zu vermeiden.

Community Services

Der Aufbau und die Förderung von Communities, sind wesentliche Elemente, um den erforderlichen kulturellen Wandel in der Wissenschaft in Bezug auf den Umgang mit Forschungssoftware zu unterstützen. Dazu soll die Etablierung von Communities of Practice¹⁷ in und zwischen den Zentren gefördert werden. Schließlich sollen die Zentren beim Aufbau von Entwickler- und Nutzer-Communities für konkrete Forschungssoftware unterstützt werden.

Ein wesentlicher Aspekt besteht im Aufbau und dem Betrieb des aktiven Austauschnetzwerkes zwischen den Helmholtz-Zentren. Das Netzwerk dient auch zur inhaltlichen Einbindung der Zentren in das Arbeitspaket und

16 Ein Beispiel für ein solches Best Practices Kennzeichnungssystem stellt das Best Practices Badge der Linux Foundation dar: <https://github.com/coreinfrastructure/best-practices-badge>.

17 Als Community of Practice bezeichnet man praxisbezogene Gemeinschaften von Personen, die vor ähnlichen Aufgaben stehen und zu deren Lösung voneinander lernen wollen: <http://wenger-trayner.com/introduction-to-communities-of-practice/>

zur Erhöhung der Nutzerakzeptanz. Dazu soll eine schlanke Koordinationsstelle aufgebaut werden, welche die Zentren effektiv einbindet und den Austausch koordiniert. Über die Bereitstellung von Webseiten, Dokumenten und vor allem durch regelmäßige direkte Kommunikation werden die Ergebnisse der Arbeitspakete in die Zentren getragen. Auf Seite der Zentren sollen dedizierte Ansprechpartner etabliert werden, die als Kommunikationspartner für die Koordinationsstelle bzw. das Software Kompetenzzentrum dienen. Zudem treiben sie die zentrums-spezifische Strategie zum nachhaltigen Umgang mit Forschungssoftware voran und werden aktiv beim Aufbau einer lokalen Entwickler-Community unterstützt.

Das Austauschnetzwerk soll genutzt werden, um den Aufbau von Communities of Practice in den Zentren sowie die Etablierung von Communities um erfolgreiche Forschungssoftware voranzutreiben. Ein praktisches Beispiel für eines solches Netzwerk ist das Software-Engineering-Netzwerk des DLR.

Training und Guidelines

Das Ziel in diesem Arbeitspaket besteht in der gezielten Aus- bzw. Weiterbildung von Mitarbeiterinnen und Mitarbeitern der Zentren im Bereich der nachhaltigen Entwicklung von Forschungssoftware. Dazu fokussiert sich der Service auf die Ausarbeitung eines übergeordneten Lernprogramms, der Konzeption und Durchführung von Trainings, Workshops, Hackathons und MOOCS sowie der Konzeption und Bereitstellung von Online-Kursen. Eine wichtige Aufgabe ist auch Ausarbeitung und Pflege von praktischen Empfehlungen zur nachhaltigen Entwicklung von Forschungssoftware.

Inhaltlich wird das Lernprogramm in einen Bereich Grundlagen und in einen Bereich für Fortgeschrittene aufgeteilt. Im Bereich der Grundlagen bietet es sich an, die Software-bezogenen Schwerpunkte gezielt mit Themen wie Publikation und Forschungsdatenmanagement zu verschränken, um den Zusammenhang aufzuzeigen. Der Bereich für Fortgeschrittene fokussiert sich insbesondere auf die Themenfelder der professionellen Software-Entwicklung. Schließlich ist anzustreben, dass die Grundlagenveranstaltungen in den Onboarding-Prozess der Zentren aufgenommen werden, so dass den Forscherinnen und Forschern bereits zu Beginn entscheidende Grundlagen vermittelt werden.

Zur Ausarbeitung und Umsetzung des Lernprogramms bietet sich eine direkte Zusammenarbeit mit der Helmholtz „Information & Data Science Academy“ (HIDA) an. Zudem besteht im Bereich der Grundlagenthemen auch die Möglichkeit mit der Organisation „The Carpentries“ zu kooperieren, dadurch können Forscherinnen und Forscher der Helmholtz-Zentren sich als Trainer ausbilden lassen. Dieser Ansatz bietet die Chance, Wissen zur nachhaltigen Software-Entwicklung stärker in den Zentren zu verankern, erlaubt es in den Zentren, leichter bedarfsorientiert Kurse selbst durchzuführen, und liefert einen Anreiz für Forscherinnen und Forscher, sich in dem Themengebiet zu engagieren. Eine solche Kooperation kann über die HIDA organisiert und koordiniert werden.

Ein weiterer Aspekt besteht in der Ausarbeitung und Pflege von Empfehlungen zur nachhaltigen Entwicklung von Forschungssoftware. Die Helmholtz-Zentren sollen die Empfehlungen als Basis für eigene Policies leicht nachnutzen können. Zudem können die Empfehlungen verwendet werden, um konkrete Mindestanforderungen an Forschungssoftware in der Helmholtz-Gemeinschaft schrittweise zu etablieren. Schließlich bilden sie die Basis, um die übrigen Software Services zielgerichtet auszuprägen. Die Empfehlungen beschreiben konkrete Praktiken und zeigen anhand von Beispielen deren Umsetzung auf. Thematisch sollen die Empfehlungen zumindest die Bereiche Entwicklungspraxis, Dokumentation, Qualitätssicherung und Publikation umfassen.

Als Ausgangspunkt für die Entwicklung der Empfehlungen bieten sich die Richt- und Leitlinien zum Umgang mit wissenschaftlicher Software der Helmholtz Open Science Initiative und der Schwerpunktinitiative Digitale Information an. Zudem existieren verschiedene Best Practice Guides in Bezug auf die nachhaltige Entwicklung von Forschungssoftware¹⁸, auf denen sich aufbauen lässt. Schließlich bietet es sich an, in diesem Bereich mit internationalen Initiativen, wie z. B. dem SSI, zu kooperieren.

Governance, Monitoring und Erfolgskontrolle

Die Entwicklung der Software Services erfolgt durch das Helmholtz Software Kompetenzcluster. In ihm sind die technischen und organisatorischen Kompetenzen und die beteiligten Zentren vertreten, die die Prozesse und Projekte organisieren und begleiten (zu weiteren Aspekten der Governance wie z. B. dem Reporting siehe Kapitel 7).

- ▶ Der Erfolg der Serviceangebote kann durch folgende Maßnahmen ermittelt werden:
- ▶ Analyse per Umfrage, Zählen von Software mit Badges (Qualitätssiegel),
- ▶ Monitoring der Anzahl der Nutzer, Softwarepakete, Aktivitäten der technologischen Services,
- ▶ Anzahl der Kurse, geschulter Mitarbeiter und Auswertung der Feedbackbögen der Weiterbildung,
- ▶ Support Services: Anfragen, Consulting, Analyse der Qualität per Fragebogen.

6 ARBEITSPAKETE UND MEILENSTEINE

Aufgrund des Umfangs der Plattformaktivitäten und der verbundenen Arbeitspakete wird ein kontinuierliches und agiles Entwicklungsmodell gewählt. Nach einer zügigen Startphase auf der Basis vorhandener Lösungen und bestehender Partnerschaften werden, der Gesamtkonzeption folgend, iterativ weitere Services ausgewählt, aufgebaut und schrittweise in die Breite der Helmholtz-Zentren überführt. Die personellen Ressourcen werden im ramp up zunächst in der Konzeptionierung, Auswahl und Entwicklung lokalisiert und mit dem Arbeitsfortschritt zunehmend für die Implementierung und Weiterentwicklung genutzt. Es wird erwartet, dass über den Einsatz des vollumfänglichen Personals bereits ab 2019 ein zügiger Start der Plattform erreicht werden kann.

Wichtig für dieses Modell sind eine sorgfältige Auswahl und Priorisierung des Portfolios von existierenden, innovativen Lösungen, ein kontinuierliches Monitoring und die Erfolgskontrolle. Mit dem Übergang von der Aufbau- und Pilotphase (Test mit ausgewählten Zentren) in die Betriebsphase erfolgen oft der schrittweise Ausbau und die Erweiterung des Nutzerkreises, mit dem Ziel, alle Zentren zu erreichen. Mit wachsendem Portfolio und damit Nutzen für die Wissenschaftler wird die Attraktivität der HIFIS Plattform steigen und so weiter Nutzergruppen anziehen, zunächst die Zentren-übergreifenden Teams.

18 Z. B. Software-Management-Pläne des SSI, Empfehlungen zur Software-Entwicklung des niederländischen eScience Center und des DLR.

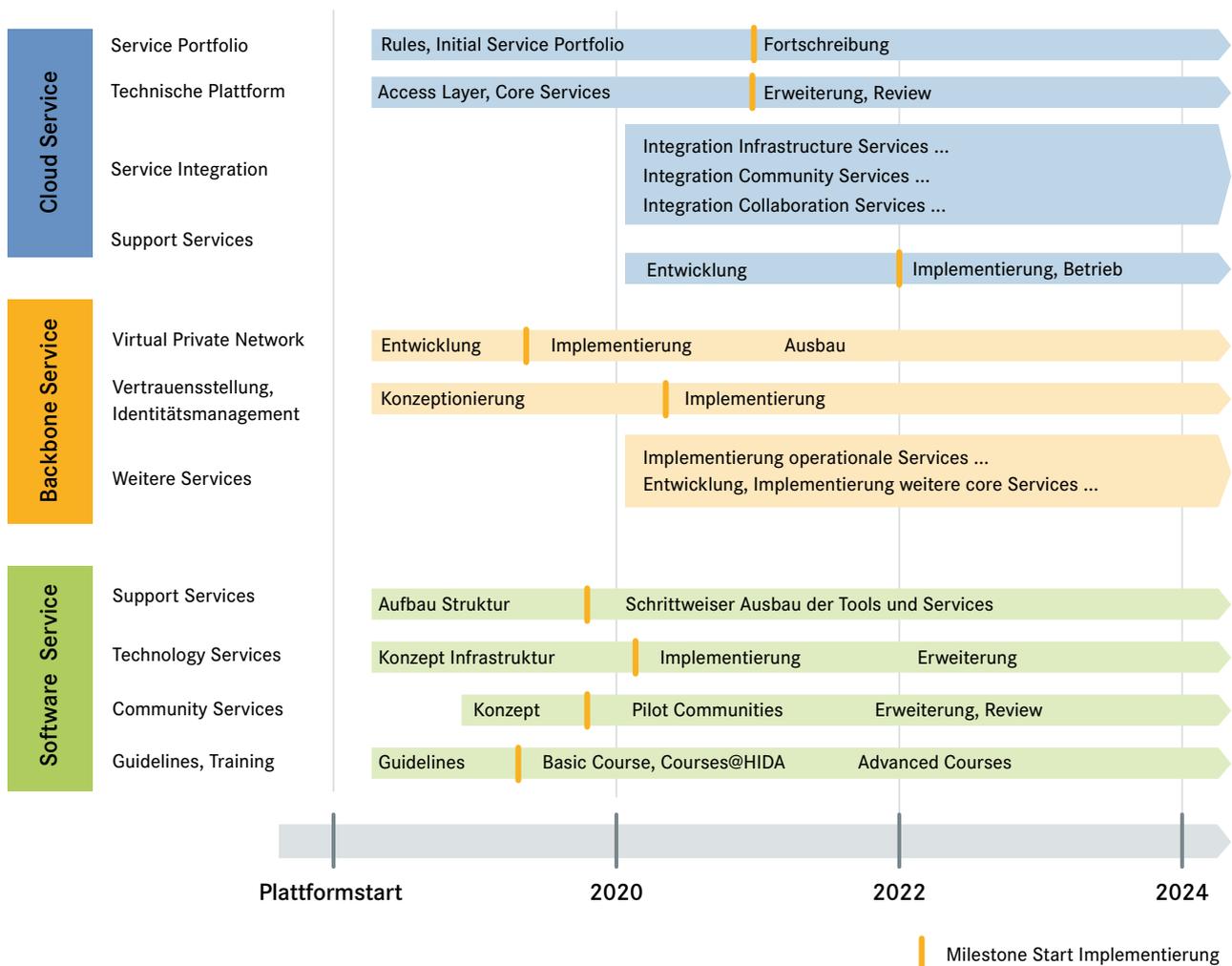


Abbildung 11: Time Chart HIFIS

7 GOVERNANCE

Die Plattform „Helmholtz Infrastructure for Federated ICT Services“ (HIFIS) soll der Stärkung und Weiterentwicklung der ICT-Kompetenzen der gesamten Gemeinschaft dienen. Das hier vorgeschlagene Vorhaben basiert daher auf einem stark förderierten Konzept, welches die vielen existierenden ICT-Kompetenzen vieler Helmholtz-Zentren verbinden kann und im Einklang mit den allgemeinen Prinzipien zur Einbettung der Plattformen in die Helmholtz-Gemeinschaft (diese werden für alle Plattformen im Einleitungskapitel beschrieben) sein soll.

Die strategischen Ziele würden von den Vorständen der Gemeinschaft festgelegt. Alle Teil-Plattformen würden mindestens einem Helmholtz-Zentrum (Host Lab) zugewiesen. Das Personal, das für HIFIS eingestellt würde, untersteht der disziplinarischen Hoheit des jeweiligen Helmholtz-Zentrums. Im Rahmen der mit der Helmholtz-Gemeinschaft vereinbarten Ziele und Leistungen bezüglich HIFIS soll das jeweilige Helmholtz-Zentrum die jeweilige Teil-Plattform operativ steuern.

HIFIS gliedert sich in die drei Kompetenzcluster Cloud Services, Backbone Services und Software Services, die eng miteinander verknüpft sind, aber jeweils eigene Kompetenzgebiete darstellen. An jedem Kompetenzcluster beteiligen sich mehrere Helmholtz-Zentren, die sich auch durch bereits vorhandene Kompetenzen qualifizieren. Jeder Kompetenzcluster soll über einen Koordinator verfügen.

Alle drei Kompetenzcluster werden durch das Helmholtz ICT Office in Verbindung gesetzt. Das Helmholtz ICT Office soll als Koordinator der gesamten HIFIS-Plattform fungieren und wird von einem Sprecher oder einer Sprecherin geleitet. Das Helmholtz ICT Office wirkt als Koordinierungsstelle und Bindeglied zwischen den Kompetenzcluster. Es stellt insbesondere auch eine technische und organisatorische Koordinierung und das Controlling und Reporting für HIFIS sicher. Es bildet auch die Schnittstelle zu vergleichbaren externen Gremien, wie sie z. B. bei der FhG oder der MPG schon existieren.

Die zentrale Koordinierung von HIFIS soll durch das Helmholtz ICT Federation Board erfolgen (dies entspricht dem Lenkungskreis im Kontext der allgemeinen Prinzipien zur Einbettung in die Helmholtz-Gemeinschaft). Entsprechend den Vorgaben der von den Vorständen der Helmholtz-Gemeinschaft vorgegebenen Ziele kann das Helmholtz ICT Federation Board den jährlichen Implementierungsplan für HIFIS aufstellen und berichtet umgekehrt über die erbrachten Leistungen an den Helmholtz-Inkubator (dieser berichtet wiederum an die Mitgliederversammlung). Das Helmholtz ICT Federation Board wird durch das Helmholtz ICT Office unterstützt. Die Zusammensetzung des Helmholtz ICT Federation Boards wird von der Mitgliederversammlung bestimmt. Es sollen neben den Plattform-Koordinatoren sowie Koordinatoren der Kompetenzcluster weitere Vertreter aus dem KODA, von der HDF und auch externe Beteiligungen vertreten sein. Der HIFIS-Koordinator und die Geschäftsstelle sollen als Gäste an den Sitzungen des Helmholtz ICT Federation Board teilnehmen können.

Es wird vorgeschlagen, HIFIS durch einen wissenschaftlichen Beirat zu flankieren, der die Entwicklung der Plattform mit den mittelfristigen Bedarfen der wissenschaftlichen Bereiche abgleicht und den Lenkungskreis und die Koordinatoren der Plattform unterstützt.

Die Mitglieder des Beirats sollen vom Präsidenten auf Vorschlag der Plattform-Koordinatoren und des Inkubators berufen werden. Die Expertinnen und Experten des Beirats sollen in regelmäßigen Abständen beraten. Sie könnten den jährlichen Plattformbericht kommentieren und mit Empfehlungen ergänzen. Dazu sollen sie regelmäßig Informationen vom Koordinator der jeweiligen Plattform erhalten.

Auf der Basis von Empfehlungen des wissenschaftlichen Beirats können ggf. zu einem späteren Zeitpunkt Folgeprojekte im IVF beantragt werden.

Der Know-how Transfer soll von den Kompetenzclustern über die lokalen Ansprechpartner zu den Forscherinnen und Forschern erfolgen. Der lokale Support erfolgt nach Bedarf Community-spezifisch unter Einbeziehung der lokalen Partner (z. B. Universitäten, lokale Research Schools) durch die Helmholtz-Zentren selbst, die dabei von den Kompetenzclustern substantiell unterstützt werden.

Die Kompetenzcluster werden jeweils von einem Koordinator und zwei Stellvertretern (Koordinatoren der Kompetenzcluster) aus unterschiedlichen Helmholtz-Zentren koordiniert und sind mit den für die Maßnahmen geplanten Stellen und finanziellen Mitteln ausgestattet. Die Binnenstruktur der Kompetenzcluster wurde jeweils im Kapitel 4 beschrieben.

Die Helmholtz-Zentren, in denen die personellen Ressourcen lokalisiert sind, sollen den Kompetenzzentren mindestens einmal jährlich über den Einsatz der zugewiesenen statischen Mittel berichten. Die Kompetenzzentren sollen wiederum an das Helmholtz ICT Office und dieses an das Helmholtz ICT Federation Board berichten.

Es ist absehbar, dass für die Helmholtz-Gemeinschaft strategische Fragen zur nationalen und internationalen Positionierung und zur ICT Versorgung entstehen, die über die hier beschriebenen Aufgaben von HIFIS hinausgehen. Aus der Arbeit von HIFIS werden übergeordnete strategische Fragen zur Weiterentwicklung der Helmholtz Daten- und ICT-Infrastruktur entstehen, die nicht innerhalb von HIFIS gelöst werden können. Nach Abstimmung mit dem wissenschaftlichen Beirat kann HIFIS das Top-Management der Helmholtz-Gemeinschaft auf solche Fragen hinweisen und auf Wunsch der Mitgliederversammlung Vorschläge unterbreiten.

8 PERSONELLE UND FINANZIELLE RESSOURCEN

Die in den Kapiteln 5 und 6 genannten Strukturen und Services werden über statische Mittel aufgebaut und betrieben. Statische Mittel sind einerseits personelle Ressourcen als auch Sachmittel. Wesentliche Beiträge für die Integration von Services, die bei den Helmholtz-Zentren vorhanden sind, werden von HIFIS-Mitarbeitern bei den Zentren geleistet. Dadurch ist besonders in der Startphase eine Realisierung von quick wins und so das schnelle Erreichen einer notwendigen Qualität zu erwarten.

Insgesamt ergibt sich folgende Planung der jährlichen Kosten aus statischen Mitteln für die Umsetzung von HIFIS. Die Werte sind Mittelwerte für den Zeitraum von 5 Jahren.

Tabelle 1: HIFIS Finanzplan: Wiedergegeben ist hier die aggregierte Kostendarstellung nach formeller Vorgabe der Geschäftsstelle der Helmholtz-Gemeinschaft. Angezeigt sind die erwarteten Kosten pro Jahr für die Plattform im voll ausgebauten Zustand, gegliedert nach Personal- und Sachkosten (grundfinanziert). Die Schätzungen der AG, die dieser Darstellung zu Grunde liegen, sind im Anhang 9.3 zu finden.

HIFIS FINANZPLAN

| Plattformanteil | Kosten p.a. in k€ |
|--|-------------------|
| Arbeitsbereich | |
| <i>Kostenart</i> | |
| Administration | |
| Koordination | |
| <i>Sachkosten</i> | 20 |
| <i>Personalkosten</i> | 340 |
| Summe | 360 |
| Cloud Services | |
| Cloud Service Portfolio | |
| <i>Sachkosten</i> | 12 |
| <i>Personalkosten</i> | 187 |
| Technische Plattform | |
| <i>Sachkosten</i> | 50 |
| <i>Personalkosten</i> | 272 |
| Services Integration | |
| <i>Sachkosten</i> | 127 |
| <i>Personalkosten</i> | 884 |
| Betrieb, Support | |
| <i>Sachkosten</i> | 14 |
| <i>Personalkosten</i> | 102 |
| Summe | 1.647 |
| Backbone Services | |
| VPN | |
| <i>Sachkosten*</i> | 223 |
| <i>Personalkosten</i> | 425 |
| Vertrauensstellung, Identitätsmanagement | |
| <i>Sachkosten</i> | 11 |
| <i>Personalkosten</i> | 255 |
| Weitere Services | |
| <i>Sachkosten</i> | 14 |
| <i>Personalkosten</i> | 255 |
| Summe | 1.183 |
| Software Services | |
| Ausbildung, Training | |
| <i>Sachkosten</i> | 21 |
| <i>Personalkosten</i> | 170 |
| Technologie Services | |
| <i>Sachkosten</i> | 64 |
| <i>Personalkosten</i> | 255 |
| Support Services | |
| <i>Sachkosten</i> | 63 |
| <i>Personalkosten</i> | 340 |
| Community Services | |
| <i>Sachkosten</i> | 11 |
| <i>Personalkosten</i> | 170 |
| Summe | 1.094 |
| Summe pro Jahr | 4.284 |
| davon: | |
| <i>Sachkosten (grundfinanziert)</i> | 629 |
| <i>Personalkosten (grundfinanziert)</i> | 3.655 |

* Darin enthalten sind 200.000 € für Network Services. Die genauen Kosten werden mit DFN verhandelt und im Laufe des Jahres 2019 ermittelt werden können. Hier erfolgt ggf. noch eine Anpassung.

9 ANHANG

9.1 ABKÜRZUNGSVERZEICHNIS

| | |
|----------|--|
| AAI | Authentication and Authorisation Infrastructure |
| AARC | Authentication and Authorisation for Research and Collaboration |
| AFAM | Institute für höhere musikalische künstlerische Ausbildung |
| AWI | Alfred-Wegener-Institut |
| BB | Backbone |
| BCC | Helmholtz Backbone Services Competence Centre |
| CC | Helmholtz Cloud Competence Centre |
| CCC | Services Competence Centre |
| CERN | European Organization for Nuclear Research |
| CNAF | Research and Development in Information and Communication Technologies |
| CNRS | Centre national de la recherche scientifique |
| DESY | Deutsches Elektronen-Synchrotron |
| DFG | Deutsche Forschungsgemeinschaft |
| DFN | Deutsches Forschungsnetz |
| DKFZ | Deutsches Krebsforschungszentrum |
| DLR | Deutsches Zentrum für Luft- und Raumfahrt |
| DOI | Digital Object Identifier |
| EGI | Federated European e-Infrastructure to provide advanced computing services |
| EMBL-EBI | European Bioinformatics Institute |
| EOSC | European Open Science Cloud |
| ESnet | Energy Sciences Network – NREN in USA |
| ESX/ESXi | Hypervisor VMware ESXi, Nachfolger von ESX |
| FAIR | Findable, Accessible, Interoperable and Reusable (Initiative) |
| FhG | Fraunhofer-Gesellschaft |
| FTS | File Transfer System |
| FZJ | Forschungszentrum Jülich |
| Gbit/s | Gigabit pro Sekunde |
| GEANT | Europäisches Datennetzwerk |
| GeRDI | Generic Research Data Infrastructure |

| | |
|--------|---|
| GFZ | Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum |
| GGUS | Global Grid User Support |
| GPU | Grafikprozessor |
| HAF | Helmholtz Analytics Framework |
| HAICU | Helmholtz Artificial Intelligence Cooperation Unit |
| HDF | Helmholtz Data Federation |
| HiBEF | Helmholtz International Beamline for Extreme Fields |
| HIDA | Helmholtz Information & Data Science Academy |
| HIDSS | Helmholtz Information & Data Science School |
| HIFIS | Helmholtz Infrastructure for Federated ICT Services |
| HIP | Helmholtz Imaging Platform |
| HMC | Helmholtz Metadata Center |
| HMGU | Helmholtz-Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt |
| HPC | High Performance Computing |
| HZB | Helmholtz-Zentrum Berlin für Materialien und Energie |
| HZDR | Helmholtz-Zentrum Dresden-Rossendorf |
| IaaS | Infrastructure as a Service |
| IAM | Identity and Access Management |
| ICT | Information and Communication Technology |
| IDM | Identity Management |
| IdP | Identity Provider |
| INFN | Nationales Institut für Kernphysik |
| IPP | Max-Planck-Institut für Plasmaphysik |
| IT | Informationstechnologie |
| JCNS | Jülicher Zentrum für Forschung mit Neutronen |
| KIT | Karlsruher Institut für Technologie |
| KNK | Kernnetzknotten |
| KODA | Koordinierungsausschuss Datenverarbeitung |
| LDAP | Lightweight Directory Access Protocol |
| LHC | Large Hadron Collider |
| LHCONE | Large Hadron Collider Open Network Environment |
| LWN | Landeswissenschaftsnetz |
| MDC | Max-Delbrück-Centrum für Molekulare Medizin in der Helmholtz-Gemeinschaft |

| | |
|-----------|--|
| MPG | Max-Planck-Gesellschaft |
| NAS | Network Attached Storage |
| NFDI | Nationalen Forschungsdateninfrastruktur |
| NIST | Nationale Institute of Standards and Technology |
| NREN | Nationale Forschungs- und Ausbildungsnetzwerke |
| NSF | National Science Foundation |
| OFFDI | Offene Föderation von Forschungsdateninfrastrukturen (Initiative) |
| OIDC | OpenID Connect |
| OLA | Operation Level Agreement |
| PaaS | Platform as a Service |
| PB | Peta Byte |
| perfSonar | Test- und Messinfrastruktur, die von Wissenschaftsnetzen und -einrichtungen zur Überwachung und Sicherstellung der Netzwerkleistung genutzt wird |
| PGI | Peter Grünberg Institut |
| QoS | Quality of Service |
| REFEDS | Research and Education FEDerations group |
| RSE | Research Software Engineers |
| SaaS | Software as a Service |
| SAML | Security Assertion Markup Language |
| SCC | Helmholtz Software Competence Centre |
| SLA | Service Level Agreement |
| SSI | Software Sustainability Institute |
| STFC | Science and Technology Facilities Council |
| SUNET | NREN in Schweden |
| SW | Software |
| SWOT | Strengths – Weaknesses – Opportunities – Treaths – Analysis |
| TU | Technische Universität |
| UFZ | Helmholtz-Zentrum für Umweltforschung |
| USB | Universal Serial Bus |
| VO | Virtuelle Organisation |
| VPN | Virtual Private Network |
| X-WiN | Glasfaser-Wissenschaftsnetz des DFN |

9.2 ÜBERSICHT ÜBER DIE PLANUNG DER AUFBAUPHASE

Die folgende Grafik stellt die zeitliche Abfolge der Arbeitspakete des im Kapitel 5 detailliert beschriebenen Arbeitsprogramms im Überblick dar.

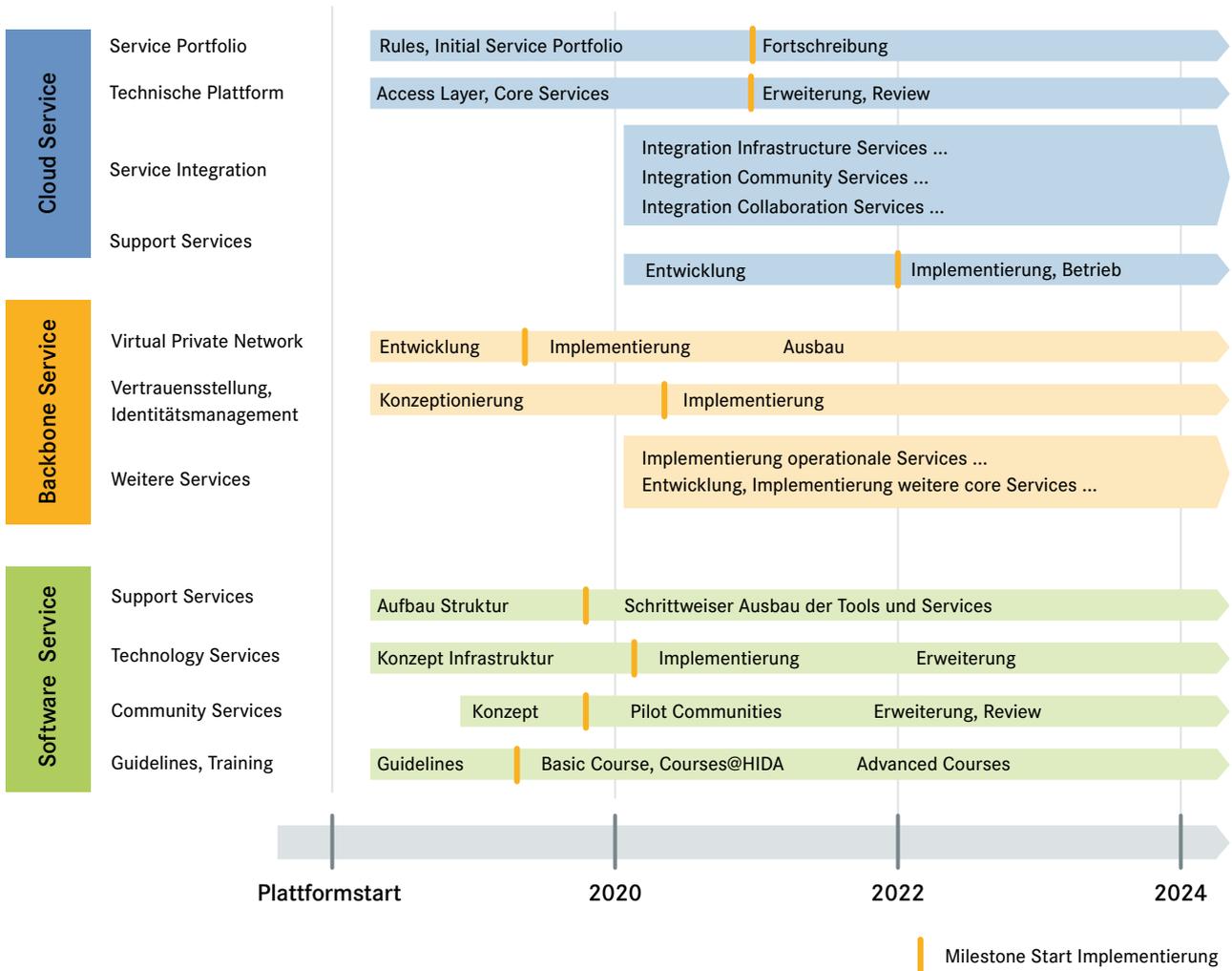


Abbildung 13: Time Chart HIFIS

9.3 HIFIS VERORTUNGSVORSCHLAG

Die folgende Übersicht zeigt die Vorschläge der Arbeitsgruppe zur Einbeziehung der Kompetenzen der Helmholtz-Zentren in die HIFIS-Plattform über die Zuordnung von FTE und die daraus resultierende Verortung der Plattformteile.

Kapitel 5 des Antrags stellt die spezifischen Arbeiten im Detail dar.

Tabelle 2: Vorschläge zur Einbeziehung der Kompetenzen der Helmholtz-Zentren in die HIFIS-Plattform über die Zuordnung von FTE (Angaben in FTE/a)

| | | Summe | AWI | DESY | DLR | DKFZ | FZJ | GFZ | HMGU | HZB | HZDR | KIT | UFZ |
|------|-------------------|-----------|----------|------------|----------|----------|----------|----------|----------|----------|------------|----------|----------|
| WP1 | Cloud Services | 17 | 1 | 3 | | 2 | 3 | | 1 | 4 | | 3 | |
| WP 2 | Backbone Services | 11 | 1 | 3 | | 1 | 2 | | 1 | | 1 | 2 | |
| WP 3 | Software Services | 11 | | | 3 | 1 | | 2 | | | 4 | | 1 |
| WP 4 | Administration | 4 | | 1,5 | | | | | | 1 | 1,5 | | |
| | Summen | 43 | 2 | 7,5 | 3 | 4 | 5 | 2 | 2 | 5 | 6,5 | 5 | 1 |

Die folgende Tabelle stellt die Zuordnung von Ressourcen zu den Unterarbeitspaketen im Überblick dar.

Tabelle 3: Zuordnung von personellen Ressourcen zu den Unterarbeitspaketen (Angaben in FTE/a)

| Arbeitspaket | | | | FTE/a |
|---------------|-------------------|--------|--|------------|
| WP1 | Cloud Services | WP 1.1 | Serviceportfolio | 2 |
| | | WP 1.2 | Technische Plattform | 3 |
| | | WP 1.3 | Services Integration | 10 |
| | | WP 1.4 | Betrieb, Support | 1 |
| | Summe | | | 17* |
| WP 2 | Backbone Services | WP 2.1 | VPN | 5 |
| | | WP 2.2 | Vertrauensstellung, Identitätsmanagement | 3 |
| | | WP 2.3 | Weitere Services | 3 |
| | Summe | | | 11 |
| WP 3 | Software Services | WP 3.1 | Ausbildung, Training | 2 |
| | | WP 3.2 | Technologie Services | 3 |
| | | WP 3.3 | Support Services | 4 |
| | | WP 3.4 | Community Services | 2 |
| | Summe | | | 11 |
| WP 4 | Administration | | | 4 |
| Summen | | | | 43 |

* Durch Rundungsdifferenzen ergibt die Summe der dargestellten FTE/a aus den Unterarbeitspaketen nicht die dargestellte Summe des WP1.

Die folgende Tabelle stellt die Verteilung der Sach- und Netzwerkkosten dar.

Tabelle 4: Sach- und Netzwerkkosten HIFIS (Angaben in €/a)

| Bereich | | Kosten |
|------------------|--|------------------|
| Sachkosten /a | Investitionen (z. B. Software) | 102.000 € |
| | Externe Zuarbeiten | 150.000 € |
| | Outreach (z. B. Veranstaltungen, Veröffentlichungen) | 40.000 € |
| | Office (2.000 €/FTE/a) | 8.000 € |
| Summe | | 300.000 € |
| Netzwerkkosten/a | | 200.000 € |
| Gesamt | | 500.000 € |

Dabei wird vorgeschlagen, die Investitionen und Kosten für externen Zuarbeiten für das Kompetenzcluster Cloud Services dem HZB und für das Kompetenzcluster Software Services dem HZDR zuzuordnen. Die Netzwerkkosten fallen im Bereich der Backbone Services an (Desy). Die Kosten für Outreach entfallen auf alle drei genannten Zentren.

Die folgende Tabelle stellt die Sach- und Netzwerkkosten sowie die Reisekosten nach Kompetenzclustern dar.

Tabelle 5: Sach- und Netzwerkkosten sowie Reisekosten nach Kompetenzclustern (Angaben in €/a)

| | Sach-/Netzwerkkosten | Reisekosten |
|-------------------|----------------------|------------------|
| Cloud Services | 151.000 € | 51.000 € |
| Backbone Services | 215.000 € | 33.000 € |
| Software Services | 126.000 € | 33.000 € |
| Administration | 8.000 € | 12.000 € |
| Gesamt | 500.000 € | 129.000 € |

Die folgende Tabelle zeigt die Gesamtkosten je Zentrum.

Tabelle 6: Gesamtkosten je Zentrum differenziert als Personalkosten, Sach- und Netzwerkkosten und Reisekosten (Angaben in €/a) gem. Verortungsvorschlag der Arbeitsgruppe

| Zentrum | Summen Personalmittel | Summen Sach-, Netzwerkkosten | Summen Reisekosten | Gesamt |
|---------------|--------------------------|---------------------------------|-----------------------|--------------------|
| AWI | 170.000 € | | 6.000 € | 176.000 € |
| DESY | 637.500 € | 218.000 €** | 22.500 € | 878.000 €** |
| DLR* | 255.000 € | | 9.000 € | 264.000 € |
| DKFZ | 340.000 € | | 12.000 € | 352.000 € |
| FZJ | 425.000 € | | 15.000 € | 440.000 € |
| GFZ | 170.000 € | | 6.000 € | 176.000 € |
| HMGU | 170.000 € | | 6.000 € | 176.000 € |
| HZB | 425.000 € | 153.000 € | 15.000 € | 593.000 € |
| HZDR | 552.500 € | 129.000 € | 19.500 € | 701.000 € |
| KIT | 425.000 € | | 15.000 € | 440.000 € |
| UFZ | 85.000 € | | 3.000 € | 88.000 € |
| Gesamt | 3.655.000 € | 500.000 € | 129.000 € | 4.284.000 € |

*Die Anteile des DLR an der Plattform müssten durch BMWi-Aufwüchse dargestellt werden.

**Darin enthalten sind 200.000€ für Network Services. Die genauen Kosten werden mit DFN verhandelt und im Laufe des Jahres 2019 ermittelt werden können. Hier erfolgt ggf. noch eine Anpassung.

9.4 UMFELDANALYSE

Wesentliche Inhalte der Analyse des Umfeldes der HIFIS-Plattform sind aufgrund der diversen Situationen für die drei Komponenten Cloud-Services, Backbone-Services und Software-Services in den Status-quo-Kapiteln des Arbeitsprogramm (Kapitel 5) dargestellt. An dieser Stelle erfolgt eine zusammenfassende Darstellung bzw. die Einbeziehung zusätzlicher Aspekte.

FACHLICHES UMFELD

Die Umsetzung von HIFIS unterstützt durch die förderierten ICT-Services insgesamt die wissenschaftliche Forschung der Helmholtz-Zentren und die Zentren-übergreifende Zusammenarbeit (sowohl innerhalb der Helmholtz-Gemeinschaft als auch mit externen Partnern) und trägt somit zum Erreichen der Ziele der Helmholtz-Gemeinschaft bei.

Helmholtz-Gemeinschaft-Intern

Die Helmholtz Data Federation weist wesentliche Schnittstellen zu den Aktivitäten der HIFIS-Plattform auf (Infrastrukturelle Hardware, Arbeiten zu AAI und Cloud Services). Während HDF vorwiegend auf Großprojekte ausgerichtet ist, ist der Ansatz von HIFIS, die Helmholtz-Gemeinschaft-Aktivitäten in der Breite unabhängig von der Größe der Aktivitäten zu unterstützen (siehe hierzu auch die Darstellungen in den jeweiligen Unterkapiteln von Kapitel 5).

Von verschiedenen Zentren werden Cloud Services angeboten. HIFIS nimmt dies auf und erstellt einen Gemeinschafts-weiten Zugang mit einer besonderen Vertrauensstellung der teilnehmenden Zentren.

Zur verfügbaren Bandbreite der Helmholtz-Zentren siehe Kapitel 5.2.

Bei einzelnen Zentren der Helmholtz-Gemeinschaft bestehen teilweise bereits VPN-Netzwerke (z. B. das voll vermaschte VPN-Netz zwischen den 20 DLR-Standorten). HIFIS zielt im Unterschied zu diesen Zentren-spezifischen Aktivitäten darauf ab, ein Helmholtz-weites Netz aufzubauen. Dabei wird es (auch im Hinblick auf Akzeptanz der Zentren) darauf ankommen, die bereits existierenden Lösungen anschlussfähig zu machen.

Verschiedene Helmholtz-Zentren arbeiten zum Thema „Nachhaltige Software“ (z. B. DLR). Im Unterschied zum Gemeinschafts-weiten Ansatz von HIFIS sind diese Arbeiten jedoch wesentlich Zentren-spezifisch ausgerichtet.

Wesentlich ist angesichts des bestehenden Umfelds die Einbeziehung der in den Zentren vorhandenen Kompetenzen.

Die Darstellungen im HIFIS-Antrag berücksichtigen die anderen Inkubator-Anträge. Dies ist z. B. für HIDA relevant, da die Software Services die Aus- und Weiterbildung als relevante Komponente enthalten, die Schnittstellen zu den HIDA-Aktivitäten enthält, aber auch zu externen Partnern (siehe hierzu auch Kapitel 9.6).

Externes Umfeld

Andere Wissenschaftsorganisationen, mit denen die Helmholtz-Zentren teilweise kooperieren aber mit denen die Forschungseinrichtungen auch teilweise in Konkurrenz um die „besten Köpfe“ stehen, haben bereits hochperformante Datennetzwerke aufgebaut bzw. realisieren dies gerade. Über HIFIS werden auf der Grundlage des bestehenden physischen Netzes z. B. des DFN wesentliche Schritte realisiert, nicht nur ein hochperformantes logisches Netzwerk zwischen den Helmholtz-Zentren zu realisieren, sondern durch die bessere Vertrauensstellung der teilnehmenden Zentren Effizienzen und Synergien zu mobilisieren.

Die in den Status-quo-Analysen (siehe Kapitel 5) dargestellten Entwicklungen z. B. bei der Zunahme der Datenmengen und der Notwendigkeit des schnellen Datenaustauschs werden als für die nächsten Jahre relevant angesehen. Eine Änderung dieser Trends, die ggf. einzelne Komponenten von HIFIS obsolet werden lassen, ist nicht erkennbar.

Forschungsgemeinschaften und die zugehörigen Netzwerke in anderen Ländern zeigen üblicherweise nicht den gleichen Zuschnitt wie die Helmholtz-Gemeinschaft im DFN. Für ein Bild der Situation in anderen Ländern wurde im Rahmen der Entwicklung dieses Antrags daher auf Forschungsverbünde fokussiert, die so wie viele Helmholtz-Zentren¹⁹ Großgeräte für wissenschaftliche Forschung nutzen. Universitäre Einrichtungen werden nur in Ausnahmefällen berücksichtigt (z. B. wenn sie im sehr engen räumlichen und organisatorischen Verbund mit den Großforschungseinrichtungen stehen).

¹⁹ siehe auch https://www.helmholtz.de/ueber_uns/helmholtz_zentren/

Das Energy Sciences Network (ESnet) in den **USA** verbindet 51 wissenschaftliche Forschungseinrichtungen und Projekte sowie 140 Forschungs- und kommerzielle Netzwerke. Es wird vom Office of Science des U.S. DoE getragen. Dem Office of Science sind 10 nationale Einrichtungen zugeordnet, von denen seit 2015 60 % mit 100 Gbit/s an das Netz angebunden sind (DOE OS 2018).

In **Schweden** sind alle Forschungs- und Bildungseinrichtungen im SUNET NREN mit einer Bandbreite von 100 Gbit/s angeschlossen (Geant Compendium 2017).

In **Italien** bildet das GARR Netzwerk den Datenverbund zwischen den Forschungszentren und Bildungseinrichtungen. Es werden sehr unterschiedliche Einrichtungen versorgt, wie z. B. Schulen, Universitäten, Bibliotheken, AFAM-Institute für höhere musikalische künstlerische Ausbildung, der Vatikan und Großforschungseinrichtungen. Fokussiert auf die Großforschungseinrichtungen des INFN (Nationales Institut für Kernphysik) mit dem CNAF (Research and Development in Information and Communication Technologies) zeigt sich ein Bild, bei dem knapp 40 % der Services eine Bandbreite im Bereich von ≥ 10 Gbit/s aufweisen (GARR 2018).

In **UK** verbindet JANET als nationales Netzwerk der Forschungs- und Bildungseinrichtungen ca. 900 Institutionen (Janet 2018). Dem Science and Technology Facilities Council (STFC) sind verschiedene Standorte zugeordnet, in denen unter anderem in den Bereichen Partikelphysik, Beschleuniger, Materie, Photon Science, Weltraumforschung, Astronomie, Biologie, Medizin, Chemie, Umwelt und Computational Science gearbeitet wird. 40 % der Standorte sind mit Geschwindigkeit von ≥ 10 Gbit/s an das Forschungsnetzwerk angebunden (Janet 2018).

Teilweise wird eine Forschungsinfrastruktur über regionale Wissenschaftsnetzwerke zur Verfügung gestellt. Zum Beispiel hat das Netzwerk des Landes Baden-Württemberg bis zum Jahr 2018 100G-Verbindungen zwischen allen Universitätsstädten in Baden-Württemberg geschaltet (BelWü 2018).

Besonders auf europäischer Ebene werden derzeit wesentliche Entwicklungen in Richtung Cloud Services und AAI/IAM im Kontext von „Open Science“ vorangetrieben. Die Berücksichtigung und teilweise aktive Begleitung dieser Entwicklungen durch HIFIS sichert die Anschlussfähigkeit der Helmholtz-Gemeinschaft an diese Entwicklungen.

Internationale Institute unterstützen die Entwicklung von nachhaltiger Software (siehe Kapitel 5.3). HIFIS sorgt in diesem Umfeld dafür, dass Helmholtz auf dieser übergeordneten Forschungsebene „vorne mitspielt“ und die wissenschaftlichen Inhalte, die mit der Software verbunden sind, auch nachhaltig sichtbar und wirkungsvoll sind.

SOZIAL

Die Helmholtz-Gemeinschaft weist für alle in HIFIS dargestellten Know-how-Felder bereits jetzt wesentliche Kompetenzträger auf. Die notwendige personelle Erweiterung für die Umsetzung der HIFIS-Inhalte kann damit auf einem sehr guten fachlichen Netzwerk erfolgen.

Die Diskussionen mit Stakeholdern der anderen Zentren im Rahmen von Projektpräsentationen und individuellen Gesprächen zeigten breite Unterstützung des Ansatzes von HIFIS. Über die Einbindung von verschiedenen Zentren in die Plattform wird eine praktische Unterstützung in der Umsetzungsphase erreicht.

ZEITLICH

Die Darstellungen zum Status quo im Kapitel 5 zeigen auf, dass parallele Entwicklungen zu den HIFIS-Services in anderen Wissenschaftsorganisationen und anderen Ländern bei der Umsetzung von HIFIS berücksichtigt werden müssen. Die Umsetzung muss z. B. so erfolgen, dass sich einzelne Zentren nicht auf Services anderer Anbieter festlegen, die eine Zusammenführung in den HIFIS-Services nur noch mit großem Aufwand realisierbar machen. Gleichzeitig können jedoch die Erfahrungen aus der Umsetzung von Projekten in HIFIS genutzt werden und ggf. zu einer wesentlich schnelleren Umsetzung führen (z. B. aus den Baden-Württembergischen und Sächsischen Erfahrungen beim Identitätsmanagement oder den EGI-Erfahrungen zu Cloud Services).

WEITERE ASPEKTE ZUM UMFELD DER SOFTWARE SERVICES

Für den Bereich der Software Services werden im Folgenden Verweise auf externe Aktivitäten, die in der Status-Quo-Darstellung im Kapitel 5 kurz angesprochen wurden, weiter ausgeführt.

Das Software Sustainability Institute (SSI)²⁰ wurde 2010 in Großbritannien gegründet. Es verfolgt das Ziel, die Nachhaltigkeit und die Qualität von Forschungssoftware in der britischen Forschungsgemeinschaft zu fördern und zu verbessern. Dazu baut es gezielt Netzwerke zwischen allen Stakeholdern der dortigen Forschungsgemeinschaft auf und fördert gezielt Aspekte wie die Weiterentwicklung von Fähigkeiten der Forscherinnen und Forscher im Bereich Software-Entwicklung, die Verbesserung der Anerkennung für das Engagement im Bereich nachhaltiger Entwicklung von Forschungssoftware sowie die Schaffung neuer Karrierewege im akademischen Umfeld für wissenschaftliche Software-Entwicklerinnen und -Entwickler. Diese Förderung übt das SSI direkt aus oder indirekt durch die Kooperation mit Organisationen wie beispielsweise der *Research Software Engineering Association*²¹ oder die *The Carpentries*²². Aktuell befindet sich ein amerikanisches Research Software Sustainability Institute²³ nach diesem Vorbild in der Planungsphase. Zudem existiert mit dem eScience Center²⁴ ein vergleichbares Institut, welches sich der Verbesserung der niederländischen Wissenschaftsexzellenz durch bessere Forschungssoftware verschrieben hat.

The Carpentries ist eine internationale Community, mit dem Ziel, Forscherinnen und Forschern essentielle Kenntnisse in den Bereichen Software-Entwicklung und Datenwissenschaften zu vermitteln. Die Community hat verschiedene Kursprogramme entwickelt, die sich an Anfänger richten und ihnen gezielt praktisch anwendbares Wissen vermitteln. Die Organisation eines Workshops kann die Community gegen eine Gebühr direkt unterstützen. Prinzipiell ist es aber möglich, auf Basis der frei verfügbaren Kursmaterialien Workshops selbstständig durchzuführen. Die Trainer sind Freiwillige und führen die Workshops unentgeltlich durch. Für deren Ausbildung hat *the Carpentries* ein eigenes Ausbildungsprogramm entwickelt. Ein wesentlicher Ursprung von *the Carpentries* ist *Software Carpentry*, welche seit 2012 bereits weltweit hunderte Veranstaltungen für mehr als 34.000 Teilnehmerinnen und Teilnehmer durchgeführt hat²⁵. Zudem haben sich bereits eine Vielzahl weiterer Carpentry Communities gebildet (z. B. HPC Carpentry, Librarians Carpentry), welche sich Anfang 2018 erstmals im Rahmen der CarpentryCon²⁶ ausgetauscht haben.

20 vgl. <https://www.software.ac.uk/about>

21 vgl. <http://rse.ac.uk/about/the-association/>

22 vgl. <https://carpentries.org/about/>

23 vgl. <http://urssi.us/>

24 vgl. <https://www.research-software.nl/about>

25 vgl. <https://software-carpentry.org/about/>

26 vgl. <http://www.carpentrycon.org/>

Die *Research Software Engineers Association* ist das Forum der britischen Research Software Engineers (RSE). Der Begriff *Research Software Engineer* bezeichnet im Forschungsbetrieb tätige Personen, welche die Bedeutung von guter Software-Entwicklungspraxis und guter wissenschaftlicher Praxis kennen und sich darum kümmern²⁷. Die RSE Association organisiert seit 2016 eine jährliche Konferenz, die mit zuletzt 224 Teilnehmerinnen und Teilnehmern zunehmend an Bedeutung gewinnt. Aufgrund des steigenden Interesses auch außerhalb Großbritanniens, haben sich eine Reihe nationaler RSE-Vereinigungen gegründet²⁸. Beispielsweise befindet sich aktuell der deutsche Ableger (de-RSE²⁹) in der offiziellen Vereinsgründung. In diesem Zusammenhang ist auch eine deutschsprachige RSE-Konferenz für das Jahr 2019 in Planung.

Im Rahmen der Open Science Initiative der Helmholtz-Gemeinschaft wurde Ende 2016 die Task Group *Zugang zu und Nachnutzung von wissenschaftlicher Software* gegründet. Die Arbeitsgruppe hat bisher ein Positionspapier³⁰ veröffentlicht und konkrete Empfehlungen³¹ an die Helmholtz-Zentren zur Schaffung von Leit- und Richtlinien zum Umgang mit wissenschaftlicher Software formuliert. Dazu zählen die Bereiche „Anreize und Metriken“, „Softwareentwicklungs- und Dokumentationspraxis“, „Zugänglichmachung, Publikations- und Transferstrategien“, „Infrastrukturen“, „Qualitätssicherung“, „Lizenzierung und weitere rechtliche Themen“ sowie „Aus- und Weiterbildung“. Aktuell erarbeitet die Task Group auf dieser Basis eine Muster-Richtlinie für die Helmholtz-Zentren.

Auch über die Grenzen der Helmholtz-Gemeinschaft hinweg gewinnt das Thema Forschungssoftware innerhalb der Deutschen Forschungsgemeinschaft an Bedeutung. Beispielsweise wurde eine Arbeitsgruppe zum Thema wissenschaftliche Software innerhalb der Schwerpunktinitiative Digitale Information gegründet. Die Arbeitsgruppe hat Anfang 2018 eine Handreichung³² veröffentlicht, die Handlungsempfehlungen an verschiedene Rollen innerhalb des Forschungsbetriebes in Bezug auf den Umgang mit Forschungssoftware ausspricht.

VERZÄHNUNG DER HIFIS-PLATTFORM MIT DEN HDF-AKTIVITÄTEN

Dieses Kapitel thematisiert die Verzahnung der HIFIS-Plattform mit den Aktivitäten der HDF im Rahmen der Inkubator-Aktivitäten in einer Gesamtperspektive. Der Inhalt dieses Kapitels ist mit dem HDF Koordinator Prof. Dr. Achim Streit im Rahmen der Antragsstellung abgestimmt worden. Einzelne Aspekte sind in den Kapiteln zur Struktur der Aktivitäten (Kapitel 4) und zum Arbeitsprogramm (Kapitel 5) dargestellt.

Die Helmholtz-Gemeinschaft stellt ihren Wissenschaftlerinnen und Wissenschaftlern und auch den wissenschaftlichen Nutzern außerhalb der Helmholtz-Gemeinschaft ein hervorragendes Angebot an wissenschaftlichen Großgeräten und Forschungsinfrastrukturen zur Verfügung. Die Nutzung dieser Einrichtungen basiert dabei zu großen Teilen auch auf Informations- und Kommunikationstechnologien, die immer stärker die Arbeits- und Forschungsprozesse beeinflussen. Die Digitalisierung verändert bereits heute die Arbeitsweisen und Methoden in den Wissenschaften und wird zukünftig die Weiterentwicklung und Neugestaltung der Forschungsarbeit prägen. Diesen Anforderungen aus der Wissenschaft müssen die IT-Infrastrukturen der Zukunft gewachsen sein, damit die Helmholtz-Gemeinschaft Lösungen für die drängenden gesellschaftlichen Herausforderungen der Zukunft entwickeln kann.

27 vgl. <http://rse.ac.uk/what-is-an-rse/>

28 vgl. <https://researchsoftware.org/>

29 vgl. <http://www.de-rse.org>

30 vgl. <https://os.helmholtz.de/index.php?id=2766>

31 vgl. <https://os.helmholtz.de/index.php?id=3197>

32 vgl. <http://doi.org/10.5281/zenodo.1172970>

HIFIS (Helmholtz Infrastructure for Federated ICT Services) ist ein Plattform-Vorschlag aus der strategischen „Inkubator“-Initiative der Helmholtz-Gemeinschaft. HIFIS zielt auf eine hervorragende Infrastruktur in der Informationstechnologie für alle Zentren der Helmholtz-Gemeinschaft ab, um die Wissenschaft auf allen Ebenen zu fördern. HIFIS unterstützt das breite Spektrum aller Forschungsarbeiten und soll aus folgenden Teilkomponenten bestehen:

- ▶ **Cloud Services:** Spezifisches Anliegen dieser Komponente ist die Bereitstellung einer föderierten Cloud-Plattform für Services, die der gesamten wissenschaftlichen Community und den Partnern in der Fläche (also nicht nur in ausgewählten Großprojekten) angeboten werden.
- ▶ **Backbone Services:** Durch diese Komponente erfolgt die Bereitstellung einer hoch performanten Netzwerkinfrastruktur zwischen allen Zentren. Die Netzwerkinfrastruktur soll insbesondere die Zentren-übergreifende Nutzung wissenschaftlicher Instrumente ermöglichen sowie eine Reihe von Basis-Services, die für kollaboratives Arbeiten notwendig sind, bereitstellen.
- ▶ **Software Services:** Über diese Komponente werden zum einen Forscherinnen und Forscher ausgebildet und unterstützt, um sie zur Entwicklung und Publikation von qualitativ angemessener, nachhaltiger Software zu befähigen. Zum anderen wird die Nachnutzung vorhandener Software durch Forscherinnen und Forscher befördert.

Die **Helmholtz-Datenföderation (HDF)** ist eine strategische Investitionsinitiative der Helmholtz-Gemeinschaft, die sich der Bewältigung der Datenflut in der Wissenschaft in ausgewählten Use Cases, insbesondere mit Bezug zu großen Forschungsinfrastrukturen, der Helmholtz-Zentren annimmt (Helmholtz 2018a). Mit der HDF wird eine föderierte Forschungsdateninfrastruktur etabliert, die konzeptionell aus drei Säulen besteht (Projektplan der HDF): 1) Software zum verteilten Daten-Management, 2) Benutzersupport mit joint R&D sowie 3) Speicher- und Analyse-Hardware. Die Realisierung der Föderation der Datenmanagementsysteme über die 6 HDF-Standorte erfolgt über die Arbeitspakete

- ▶ AP1 - Authentication and Authorization Infrastructure (AAI)
- ▶ AP2 - Online-Daten-Spiegel
- ▶ AP3 - Datentransferservice
- ▶ AP4 - Verteiltes Archiv
- ▶ AP5 - Hilfsdienste (GGUS, Portal/Wiki, Webseite, Domain)
- ▶ AP6 - Föderierte Cloud
- ▶ AP7 - Föderierte Suche nach Daten
- ▶ AP8 - Helmholtz-VPN-Ring

Eine enge **Verzahnung der HIFIS-Plattform mit den Aktivitäten der HDF** wird sichergestellt. Schnittstellen werden sich nach der bisherigen Planung vor allem in den Bereichen AAI/IAM und Cloud Services ergeben. So hat die HDF beispielsweise im Arbeitspaket 1 „Authentication and Authorisation Infrastructure (AAI)“ zum Zugang föderierter Nutzer zu föderierten Diensten an Standorten von HDF-Zentren mit dem lokalen Konto der Benutzer gearbeitet sowie im Arbeitspaket 6 „Föderierte Cloud“ zum föderierten Zugang zu HDF-Daten und -Diensten. Insbesondere beim Thema AAI/IAM können die Arbeiten im HDF-AP1, welches wiederum auf Arbeiten in europäischen Projekten fußt, als ideale Grundlage für eine Helmholtz-weite AAI/IAM-Lösung mit voller Kompatibilität zur nationalen DFN-AAI und zu europäischen AAI/IAM-Lösungen im Kontext GEANT/EduGain und EOSC dienen.

In der Folge der Abstimmung über Arbeitsinhalte zwischen HDF und HIFIS im Verlauf der HIFIS-Antragsentwicklung wurde festgelegt, dass die Arbeiten des HDF-AP8 (Helmholtz-VPN-Ring) bis auf einen Prototypen am AWI an HIFIS ausgelagert werden.

Die gewünschte enge Verzahnung wird insbesondere auch über die sowohl an HDF als auch an HIFIS beteiligten Personen gesichert. Weiterhin gibt es engste personelle Überschneidungen zwischen HDF und HIFIS, insbesondere durch Volker Gülzow (DESY) als einen der HIFIS-Berichterstatter und gleichzeitiges Mitglied des HDF Executive Board (EB).

Über diese Vorgehensweise bzw. Verzahnung ist zum einen sichergestellt, dass vorhandenes Know-how in den Zentren genutzt wird. Zum anderen kann auf diesem Weg an die bereits geleisteten Arbeiten angeknüpft, diese effizient ergänzt, Interoperabilität hergestellt und Doppelentwicklungen vermieden werden.

9.5 SWOT-ANALYSE

In den Kapiteln 5.1 bis 5.3 wurden die einzelnen HIFIS-Komponenten im Detail dargestellt. Die folgende Übersicht wertet wesentliche Eckpunkte in Form einer SWOT-Analyse (Strengths – Weaknesses – Opportunities – Threats) aus:

- ▶ „Stärken“ beschreiben in dieser Analyse, warum die Realisierung der HIFIS-Plattform in der Helmholtz-Gemeinschaft besonders sinnvoll ist und warum die Realisierung als aussichtsreich angesehen wird,
- ▶ „Schwächen“ beschreiben, welche Problemstellungen in der Helmholtz-Gemeinschaft im Zuge der Realisierung der HIFIS-Plattform angegangen werden müssen,
- ▶ „Chancen“ zeigen auf, wie die Forschung in der Helmholtz-Gemeinschaft durch die Realisierung der HIFIS-Plattform verbessert werden kann und welche Möglichkeiten hierdurch realisiert werden können (Besserstellung im Vergleich zu anderen Forschungsgemeinschaften),
- ▶ unter „Risiken“ werden Aspekte betrachtet, die die Realisierung der HIFIS-Plattform gefährden könnten.

Im Anschluss erfolgt eine Darstellung der Potenziale zum Umgang mit den benannten Schwächen und Risiken.

STÄRKEN

- ▶ Der bei der Helmholtz-Gemeinschaft vorhandene qualitativ hochwertige „Primärdatenschatz“, dessen Umfang zukünftig weiter zunehmen wird, erzeugt einen starken Bedarf für ein hoch performantes Netz und angepasste Cloud Services.
- ▶ Es besteht bereits jetzt beim DFN ein hochwertiges physisches Netzwerk, durch dessen optimierte Nutzung ein wesentlicher Schritt für eine Optimierung der Forschung bei zu erwartenden vergleichbar geringen Kosten (siehe auch: Science|Business Network's Cloud Consultation Group (2018)) getan werden kann.
- ▶ In den Zentren der Helmholtz-Gemeinschaft existieren bereits Cloud-Lösungen, die durch die Föderation besser bzw. breiter nutzbar gemacht werden.
- ▶ Die Elemente in der HIFIS-Plattform basieren überwiegend auf vorhandenen, innovativen Best-Practice Services, die Gemeinschafts-weit zugänglich gemacht werden.

- ▶ Der Nutzen der HIFIS-Plattform kommt potenziell allen Zentren zugute. Es wird ein ganzheitliches Servicekonzept bereitgestellt, das durch verbindende Workflows Mehrwerte für die Zentren und für die Gemeinschaft erbringt. Die Ergebnisse und Rückmeldungen aus Workshops und Gesprächen haben gezeigt, dass sowohl bei der Helmholtz-Geschäftsstelle als auch in den Helmholtz-Zentren eine Unterstützung des HIFIS-Ansatzes besteht.
- ▶ Es existiert ebenfalls ein großer Schatz an Forschungssoftware, dessen Umfang zukünftig weiter zunehmen wird. Es entsteht somit ein steigender Bedarf an praktischen Angeboten zur Förderung der nachhaltigen Software-Entwicklung, um die Nachvollziehbarkeit und Reproduzierbarkeit von wissenschaftlichen Ergebnissen nicht zu gefährden.

CHANCEN

- ▶ Es wird ein qualitativer Sprung in der Forschung, ihrer Transdisziplinarität und ihrer Zentren-übergreifenden Ausrichtung sowie der Sichtbarkeit der Helmholtz-Gemeinschaft durch die Realisierung der HIFIS-Plattform erreicht.
- ▶ Im Vergleich zu individuellen Lösungen wird durch den ganzheitlichen Gemeinschaft-weiten Ansatz der HIFIS-Plattform ein Kostenvorteil erwartet.
- ▶ Die Helmholtz-Gemeinschaft zeigt Referenzen als kompetenter Anbieter für Zentren-übergreifende Cloud-Angebote an die Wissenschaft.
- ▶ Die Plattform stellt Schnittstellen bereit, um die Helmholtz-Gemeinschaft mit den nationalen und internationalen Initiativen zu vernetzen (organisatorisch und technisch).
- ▶ Durch HIFIS werden Arbeitsbedingungen geschaffen, die der Arbeit in der Helmholtz-Gemeinschaft für die besten Köpfe eine zusätzliche Anziehungskraft geben und sie effektiver macht.
- ▶ Die Realisierung der Plattform trägt dazu bei, wesentliche Teile der Forschung „unabhängiger von den Einschränkungen der Geografie“ zu machen (durch entferntes Arbeiten und einen optimierten Datenfluss).
- ▶ Durch optimierte Datenhaltung und verbesserten Datenfluss wird das Risiko minimiert, Daten zu verlieren (sei es durch Mangel an Speicherplatz, der die Forscher zur Löschung von Daten zwingt (siehe hierzu auch die Ausführungen im HIP-Konzept) oder durch risikobehaftete Speichermedien (siehe hierzu auch use case Erde und Umwelt)).
- ▶ Die Thematik „nachhaltige Entwicklung von Forschungssoftware“ wird durch die HIFIS-Plattform substantiell auf praktischer Ebene vorangebracht und die positive Sichtbarkeit der Helmholtz-Gemeinschaft auch bei diesem wichtigen Thema erhöht.

SCHWÄCHEN

- ▶ Aktuell stellt sich die Netzwerk- und Dateninfrastruktur heterogen dar mit einer Vielzahl Zentren-spezifischer Lösungen und oftmals einer Zentrenorientierung der Strukturen.
- ▶ Die IT-Infrastruktur hinkt dem Trend der zunehmenden Bedeutung der Zentren- und bereichsübergreifenden Forschung hinterher.

- ▶ Die Zentren müssen sich aktiv einbringen und gemeinsame Lösungen schaffen und nutzen.
- ▶ Bei den Zentren besteht teilweise eine schwach ausgeprägte Anerkennung der Bedeutung von Forschungssoftware für die Nachvollziehbarkeit und Nachhaltigkeit der erzielten wissenschaftlichen Erkenntnisse.

RISIKEN

- ▶ Zur Realisierung der Plattform wird es notwendig sein, ausreichend Fachpersonal zu finden und für den dauerhaften Betrieb zu binden.
- ▶ Um die Bereitschaft der Beteiligung der Zentren sicherzustellen, muss die Realisierung der Plattform erfolgen, bevor Zentren in wesentlichem Umfang andere Dienste in Anspruch nehmen, in denen ihre Daten oder Prozesse „gefangen“ werden.
- ▶ Die Nutzer nehmen das Angebot der Services nicht an.
- ▶ Die Strukturen, die durch HIFIS entwickelt werden sind nicht kompatibel mit z. B. europäischen bzw. globalen Standards. Relevante User nutzen nicht-kompatible Standards.

BEWERTUNG

Ein Vergleich der Stärken, Schwächen, Chancen und Risiken zeigt das insgesamt hohe Potenzial der Verbesserung der Forschung in der Helmholtz-Gemeinschaft durch die Realisierung der HIFIS-Plattform bei vergleichsweise akzeptablen bzw. handhabbaren Risiken. Die dargestellten Stärken der Helmholtz-Gemeinschaft und Chancen aus der Realisierung der HIFIS-Plattform werden als ausreichendes Gegengewicht gesehen, die Risiken zu balancieren. Zudem sind Aktivitäten vorgesehen, um die Risiken zu handhaben:

- ▶ Um in einem akzeptablen Zeitraum das Fachpersonal für den Aufbau der Services zu finden kann zusätzlich zu den üblichen Wegen (Ausschreibungen etc.) auf das sehr umfangreiche Netzwerk der Fachleute aus den Helmholtz-Zentren zurückgegriffen werden.
- ▶ Für die zeitgerechte Realisierung der Plattform wurden in dem Projektantrag ausreichende personelle Ressourcen vorgesehen.
- ▶ Die positive Rückmeldung der Zentren bei der Vorstellung der HIFIS-Plattform zeigt, dass eine nicht-Aannahme der Services sehr unwahrscheinlich ist.
- ▶ Um das Risiko der nicht-kompatiblen Strukturen auszuschließen erfolgt in der Startphase eine zusätzliche vertiefte Analyse der internationalen Situation und relevante Entwicklungen werden verfolgt bzw. auch teilweise aktiv begleitet. Helmholtz-Zentren, die aktiv an HIFIS mitarbeiten sind auch in relevanten internationalen Projekten zu Kernthemen involviert.

Auch die dargestellten Schwächen können durch den im Antrag dargestellten Ansatz ausgeglichen oder zumindest abgemildert werden:

- ▶ Das Konzept der HIFIS-Plattform z. B. mit seinem Access Layer zur Zugänglichmachung der verteilten Cloud-Dienste und dem globalen AAI/IAM ist geeignet, heterogene Zentrenlösungen zusammen zu bringen.
- ▶ Der HIFIS-Ansatz fördert durch seine IT-Struktur die Zentren- und bereichsübergreifende Forschung.

- Aufgrund des positiven Feedbacks der Zentren bei der Vorstellung der HIFIS-Plattform kann davon ausgegangen werden, dass das Risiko, dass die Zentren sich nicht aktiv einbringen und gemeinsame Lösungen schaffen und nutzen, handelbar ist.

9.6 ZUSAMMENARBEIT MIT ANDEREN PLATTFORMEN UND EXTERNEN PARTNERN

Neben den dargestellten Nutzen für die Forschung in der Helmholtz-Gemeinschaft ergeben sich spezifische Schnittstellen zu den anderen Inkubator-Plattformen. Einzelaspekte der jeweiligen Plattformkomponenten (Cloud Services, Backbone Services, Software Services) sind in den Kapiteln 5.1 bis 5.3 dargestellt.

Generell ermöglicht bzw. optimiert HIFIS für HIP, HIDA, HAICU und HMC die Zugänglichkeit von Instrumenten der Kollaboration und von Cloud-Services, indem die notwendigen Infrastrukturen (Backbone Services, Cloud Services, Vertrauensstellungen) bereitgestellt werden. Hierdurch wird beispielsweise die Bearbeitung großer Datenmengen wie z. B. in HIP und HAICU (z. B. Datentransfer und Datenzugang) sowie die entfernte Datenbearbeitung und entfernte Zusammenarbeiten z. B. im Rahmen von HIDA unterstützt. Ein Beispiel für das Zusammenspiel von Lehre und Cloud Services ist das bwLehrpool, das eine landesweite zentrale Infrastruktur für gemeinsame virtuelle Lehrpools³³ für alle Hochschulen des Landes Baden-Württemberg entwickelt (Wilson 2013).

HMC kann für HIFIS wichtigen Input für die Auffindbarkeit relevanter Software (aber auch Cloud-Services) in Repositories bringen. Umgekehrt stellt eine über HIFIS etablierte AAI eine wichtige Grundlage für die Umsetzung der HMC-Arbeiten dar.

Die Software Services und die Unterstützung bei der Erstellung qualitativ hochwertiger Software bilden eine weitere wichtige Schnittstelle zu den Aktivitäten der anderen Plattformen, speziell für die Erstellung von Software in den wissenschaftsbasierten Plattformen HIP und HAICU. Für die Aktivitäten von HIDA resultieren Grundlagen und Materialien für Ausbildung und Training aus den Software Services.

Umgekehrt fördert HIDA die Ausbildung qualifizierten Nachwuchses für die HIFIS Plattform.

9.7 BIBLIOGRAFIE

AARC (2018): AARC: Authentication and Authorisation for Research and Collaboration, <https://aarc-project.eu/>, Letzter Zugriff 30.07.2018

Allianz (2018): Allianz der deutschen Wissenschaftsorganisationen: Schwerpunktinitiative Digitale Information, <https://www.allianzinitiative.de/>, Letzter Zugriff: 31.05.2018

BaWü (2018): bwIDM - Förderiertes Identitätsmanagement der Baden-Württembergischen Hochschulen, <https://www.bwidm.de/>, Letzter Zugriff 15.05.2018

Bell, G. (2015): The Energy Sciences Network: Overview, Update, Impact, ASCAC Meeting, American Geophysical Union, Washington, DC, March 24, 2015, Gregory Bell, Ph.D., https://science.energy.gov/~media/ascr/ascac/pdf/meetings/20150324/Bell_ESNet.pdf, Letzter Zugriff 15.05.2018

33 siehe hierzu auch <https://www.bwlehrpool.de/doku.php>

- BelWÜ (2018): BelWü - das Landeshochschulnetz, <https://www.belwue.de/>, Letzter Zugriff 15.05.2018
- BMFSFJ (2016): BMFSFJ (HG): Digitalisierung – Chancen und Herausforderungen für die partnerschaftliche Vereinbarkeit von Familie und Beruf, Berlin, 2016, <https://www.bmfsfj.de/bmfsfj/service/publikationen/digitalisierung---chancen-und-herausforderungen-fuer-die-partnerschaftliche-vereinbarkeit-von-familie-und-beruf/109006>, Letzter Zugriff 15.07.2018
- BWIDM (2018): BWIDM: Föderiertes Identitätsmanagement in Baden Württembergischen Wissenschaftseinrichtungen, <https://www.bwidm.de/>, Letzter Zugriff 30.07.2018
- DFN (2017): pers. com. DFN
- DFN (2018): pers. com. DFN
- DFN (2018a): DFN: DFN-AAI - Authentifikations- und Autorisierungs-Infrastruktur, <https://www.dfn.de/dienstleistungen/dfnaai/>, Letzter Zugriff 30.07.2018
- DFN (2018b) DFN: DFN-VPN - Große Bandbreiten zwischen festen Zielen, <https://www.dfn.de/dienstleistungen/dfnvpn/>, Letzter Zugriff 30.07.2018
- DOE OS (2018): <https://www.es.net/engineering-services/the-network/connected-sites/>, Letzter Zugriff 15.07.2018
- EC (2018): European Commission DG Research and Innovation: EOSC Strategic Implementation Roadmap 2018, Brüssel, 2018 http://ec.europa.eu/research/openscience/pdf/eosc_strategic_implementation_roadmap_large.pdf#view=fit&pagemode=none), Letzter Zugriff 30.06.2018
- EOSC (2018): European Commission: European Open Science Cloud (EOSC), <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>, Letzter Zugriff 30.07.2018
- EOSC (2018a): EOSCPilot, <https://eoscpiilot.eu/>, Letzter Zugriff 30.07.2018
- EOSC HUB (2018): Services for the European Open Science Cloud, <https://www.eosc-hub.eu/>, Letzter Zugriff 15.07.2018
- ESNet (2015): The Energy Sciences Network: Overview, Update, Impact, ASCAC Meeting, American Geophysical Union, Washington, DC ,March 24, 2015, Gregory Bell, Ph.D., https://science.energy.gov/~media/ascr/ascac/pdf/meetings/20150324/Bell_ESNet.pdf, Letzter Zugriff 11.04.2018
- GARR (2018): <https://www.garr.it/en/infrastructures/network-infrastructure/connectivity-services?key=all>, Letzter Zugriff 11.04.2018
- GARR (2018a): https://www.garr.it/images/2018-01_mappaGARR_Net.png, Letzter Zugriff 11.04.2018
- Geant Compendium (2016): GÉANT Compendium of National Research and Education Networks in Europe, 2016 Edition, <https://compendium.geant.org>, Letzter Zugriff 15.05.2018
- GEOSS (2018): Global Earth Observation System of Systems (GEOSS), <http://www.geoportal.org/>, Letzter Zugriff 30.07.2018
- Helmholtz (2013): Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V.: Open access Richtlinie des Impuls- und Vernetzungsfonds, Berlin, 2013, <https://os.helmholtz.de/open-science-in-der-helmholtz-gemeinschaft/open-access-richtlinien/open-access-richtlinie-des-impuls-und-vernetzungs-fonds-2013/>, Letzter Zugriff 15.07.2018

- Helmholtz (2016): Helmholtz-Gemeinschaft der Forschungszentren: Open Science in der Helmholtz-Gemeinschaft, <https://os.helmholtz.de/open-science-in-der-helmholtz-gemeinschaft/>, Letzter Zugriff 30.07.2018
- Helmholtz (2016): Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V.: Open access Richtlinie der Helmholtz-Gemeinschaft, <https://os.helmholtz.de/open-science-in-der-helmholtz-gemeinschaft/open-access-richtlinien/open-access-richtlinie-der-helmholtz-gemeinschaft-2016/>, Letzter Zugriff 15.07.2018
- Helmholtz (2018): Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V. (HG): Die Strategie der Helmholtz-Gemeinschaft, Berlin, 2018, https://www.helmholtz.de/fileadmin/user_upload/03_ueber_uns/Helmholtz_Strategiebrochure.pdf, Letzter Zugriff 15.07.2018
- Howinson (2016): James Howison and Julia Bullard. 2016. Software in the scientific literature: Problems with seeing, finding, and using software mentioned in the biology literature. *Journal of the Association for Information Science and Technology* 67, 9 (2016), 2137–2155. <https://doi.org/10.1002/asi.23538>, Letzter Zugriff 15.07.2018
- INDIGO (2018): INDIGO Data Cloud - Better software for better science, <https://www.indigo-datacloud.eu/>, Letzter Zugriff 30.07.2018
- Janet (2018): JANET: About the Janet Network; <https://www.jisc.ac.uk/janet>, Letzter Zugriff 15.07.2018
- Marletta, M. (2018): e-mail conversation with M. Marletta (GARR) at April 11th 2018
- MoE (2016): French Ministry of Education, Higher Education and Research: FRENCH NATIONAL STRATEGY ON RESEARCH INFRASTRUCTURES 2016 EDITION, https://cache.media.enseignementsup-recherche.gouv.fr/file/Infrastructures_de_recherche/16/4/infrastructures_UK_web_615164.pdf, Letzter Zugriff 11.04.2018
- MOSES (2018): Modular Observation Solutions for Earth Systems (MOSES), <https://moses.eskp.de/home/>, Letzter Zugriff 11.07.2018
- OFFDI (2018): Offene Föderation von Forschungsdateninfrastrukturen, Bericht zur Projektinitiative, unveröffentlicht
- Renater (2015): Renater 2015 - The National Telecommunication Network for Technology, Teaching and Research, https://www.renater.fr/IMG/pdf/crea_renater_ra_2014-2910c-pap.pdf, Letzter Zugriff 11.04.2018
- Science|Business Network's Cloud Consultation Group (2018): The European Open Science Cloud: Who pays for what?, Brüssel, 2018, www.sciencebusiness.net, Letzter Zugriff 11.04.2018
- Wilson (2013): Wilson, M.: Projekt bwLehrpool, Vortrag am 20.12.2013 bwIDM Abschlussveranstaltung KIT, SCC, <https://www.bwidm.de/wp-content/uploads/p25-wilson.pdf>, Letzter Zugriff 13.07.2018
- Yahyapour, R.; Wieder, P. (2013): Cloud-Dienste für die Wissenschaft https://www.mpg.de/6957705/JB_2013, Letzter Zugriff 12.07.2018

DANKSAGUNG

Wir bedanken uns bei allen Expertinnen und Experten aus der Helmholtz-Gemeinschaft die bisher am Inkubator-Prozess mitgewirkt haben und durch den Einsatz das vorliegende Dokument ermöglicht haben. Insbesondere möchten wir den Teilnehmerinnen und Teilnehmern der Arbeitsgruppe Basistechnologien und grundlegende Dienste danken:

- Sabine Attinger (UFZ)
- Achim Basermann (DLR)
- Dieter Beule (MDC)
- Gisbert Breitbach (HZG)
- Wolfgang zu Castell (HMGU)
- Ants Finke (HZB)
- Stephan Frickenhaus (AWI)
- Bernadette Fritzsch (AWI)
- Tobias Frust (HZDR)
- Jens Greinert (GEOMAR)
- Volker Gülzow (DESY)
- Holger Haas (DKFZ)
- Martin Hammitzsch (GFZ)
- Andreas Heiss (KIT)
- Guido Juckeland (HZDR)
- Werner Köckeritz (GFZ)
- Thorsten Kollegger (GSI)
- Uwe Konrad (HZDR)
- Ronny Kopischke (GFZ)
- Norbert Kroll (DLR)
- Sören Lorenz (GEOMAR)
- Daniel Mallmann (FZJ)
- Lars Mehwald (Geschäftsstelle)
- Martin Nußbaumer (KIT)
- Kars Ohrenberg (DESY)
- Knut Sander (DESY)
- Klaus Scheibenberger (KIT)
- Tobias Schlauch (DLR)
- Thomas Schnicke (UFZ)
- Achim Streit (KIT)
- Claas Teichmann (HZG)
- Alf Wachsmann (MDC)
- Martin Wimmer (DZNE)

www.helmholtz.de

