

## 9 Voraussetzungen und Herausforderungen

In der Schnittmenge von adressierten Datenräumen, beteiligten Personen (Anforderungsbeitragende) sowie einzusetzenden Informationstechnologien ergaben sich mehrere spezifische Voraussetzungen und Herausforderungen, die nachfolgend ausgeführt werden. Dabei wird auf Datenräume (Kap. 9.1), Anforderungsbeitragende (Kap. 9.2) sowie die Problematik einer vollständigen Vorabdefinition der Lösung (Kap. 9.3) eingegangen.

### 9.1 Datenräume

#### 9.1.1 CERA-2

CERA-2 wurde zur Beschreibung extrem heterogener Daten entwickelt und stellt eine entsprechend komplexe Metadatenbank dar – so setzt sich bereits der CERA-Kern aus über 40 unterschiedlichen Datenbanktabellen mit jeweils einer Vielzahl von Attributen zusammen. Da in dieser Metadatenbank sehr unterschiedlicher Daten dokumentiert werden, die einer sehr breit gefächerten, multidisziplinären Anwendergemeinde mit jeweils unterschiedlichen Bedürfnissen erschlossen werden sollten, waren zur Erreichung dieses Zieles eine Vielzahl von Anforderungen zu berücksichtigen und geeignet ineinander abzubilden. Insbesondere war es erforderlich, das komplexe Datenmodell dieser Metadatenbank den Anwendern so zugänglich zu machen, dass diesen eine schnelle und intuitive - und dabei zugleich individuelle und flexible - Datenauswahl und Auswertung ermöglicht wird.

#### 9.1.2 Zeitreihenmetadaten und punktverortete Zeitreihen

Sowohl Zeitreihenmetadaten wie punktverortete Zeitreihen erforderten die Entwicklung geeigneter Formen des Zugriffs, die es einer Vielzahl von Anwendern aus unterschiedlichen Disziplinen erlauben, diese komplexen Datenressourcen entsprechend individueller Bedürfnisse auszuwerten. Anders als bei CERA-2 kam hier deutlich erschwerend hinzu, dass weder die Gesamtheit der Zeitreihenmetadaten noch der punktverorteten Zeitreihen im Institut an zentraler Stelle oder in homogener Form vorlagen. Die jeweiligen Datenräume verteilten sich vielmehr auf mehrere, getrennt aufgebaute und heterogene Metadatenbanken und Datenbanken mit jeweils forschungsspezifischen thematischen und räumlichen Ausschnitten. Entsprechend mussten zusätzlich geeignete Wege gefunden werden, um die gegebene Heterogenität, die einer effizienten übergreifenden Erschließung von Zeitreihenmetadaten und punktverorteten Zeitreihen entgegenstand, zu überwinden.

### 9.2 Anforderungsbeitragende

#### 9.2.1 Immanente Vielschichtigkeit und Heterogenität

Für die zu entwickelnde Schnittstelle waren die vielfältigen Anforderungen einer großen Zahl unterschiedlicher Personen zu berücksichtigen und in geeigneter Weise umzusetzen. Die Anforderungsbeitragenden können dabei in zwei unterschiedliche, jeweils heterogene Akteursgruppen aufgeteilt werden, die im Folgenden als *Datennutzer* sowie *Datenbereiter* bezeichnet werden.

- **Datennutzer**

Als *Datennutzer* der zu entwickelnden Schnittstelle werden hier diejenigen Personen bezeichnet, die auf die zugänglich zu machenden Metadaten und Daten zugreifen sollen. Hierbei handelt es sich um eine heterogene, multidisziplinäre, räumlich verteilte und zeitlich veränderliche Gruppe von Wissenschaftlern aus diversen Fachgebieten, Abteilungen und Projekten mit einer großen Vielfalt an Muttersprachen, Forschungsinteressen und Anforderungen an die Datenversorgung. Individuelle Datennutzer unterscheiden sich zudem in der Ausprägung der jeweiligen Kenntnisse bspw. bezüglich Computertechnologie, Anwen-

dungssoftware, Datenräumen, Datenmodellen, Datenbanken etc. ebenso wie in ihrer jeweiligen Bereitschaft, Zeit und Aufwand in Erlernung und Anwendung neuer Werkzeuge zu investieren.

#### ▪ **Datenbereitsteller**

Als *Datenbereitsteller* der zu entwickelnden Schnittstelle werden hier diejenigen Personen bezeichnet, die die zugänglich zu machenden Metadaten und Daten administrieren. Hierbei handelt es sich ebenfalls um eine heterogene, multidisziplinäre und zeitlich veränderliche Gruppe von Wissenschaftlern aus diversen Fachgebieten, Abteilungen und Projekten mit jeweils individuellen Forschungsinteressen. Bei dieser Gruppe kann insbesondere von vertieften Kenntnissen über alle wesentlichen Details der jeweils von ihnen administrierten Daten ausgegangen werden. Personen aus der Gruppe der Datenbereitsteller können dabei in Personalunion als Datennutzer fungieren, indem sie die zu entwickelnde Lösung nutzen, um über diese auf eigene Daten und / oder Daten anderer Datenbereitsteller zuzugreifen.

### **9.2.2 Notwendigkeit einer übergreifenden Akzeptanz**

Ein Erfolg der zu entwickelnden Lösung setzte wesentlich und unabdingbar ihre Akzeptanz durch eine Vielzahl individueller Akteure voraus. Da im gegebenen Kontext weder die Bereitstellung von Daten noch eine individuelle Nutzung zur Datenversorgung „verordnet“ werden können, war es erforderlich, eine übergreifende Akzeptanz sowohl von Datennutzern wie Datenbereitstellern zu erreichen. Eine Schnittstelle, die von der Gruppe der *Datennutzer* als nicht geeignet zur Umsetzung ihrer Anforderungen empfunden und daher nicht zum Datenzugriff eingesetzt wird, hat das anvisierte Ziel verfehlt. Auf der anderen Seite war es unabdingbar, möglichst viele *Datenbereitsteller* zur Mitarbeit zu gewinnen, um den Datennutzern überhaupt einen hinreichend substantiellen Datenfundus anbieten zu können.

Dabei ist von einer wechselseitigen Beziehung zwischen der zu entwickelnden Schnittstelle und den beteiligten Akteursgruppen auszugehen. Zum einen wird eine solche Lösung für Datennutzer umso attraktiver, desto mehr Daten sie in funktionaler und effizienter Weise zugänglich macht; zum anderen steigt ihre Attraktivität für die Datenbereitsteller, wenn sie es ihnen ermöglicht, ihre jeweiligen Datenräume einer zunehmenden Zahl von Datennutzern zu erschließen. Eine umfassende Datengrundlage wiederum erlaubt die Bereitstellung weitergehender Funktionalität, die ihrerseits zu erhöhter Attraktivität für Datennutzer wie Datenbereitsteller beitragen kann.

### **9.2.3 Notwendigkeit einer Integration in den laufenden Forschungsbetrieb**

Bei der Entwicklung der angestrebten Lösung war zu berücksichtigen, dass einer Aufrechterhaltung des laufenden Forschungsbetriebes oberste Priorität sowohl für das Institut wie für Datennutzer und Datenbereitsteller zukam. Datennutzer wie Datenbereitsteller arbeiten an jeweiligen wissenschaftlichen Aufgabestellungen, die nicht zurückgestellt oder unterbrochen werden konnten. Hinzu kommt, dass sich am Institut eine internationale Wissenschaftlergemeinschaft jeweils für eine limitierte Zeit zu veränderlichen Konstellationen zusammenfindet, so dass das Erreichen vorgegebener Forschungsziele innerhalb eng umrissener Zeiträume höchsten Stellenwert für die beteiligten Akteure einnimmt.

Aus diesem Grund war auch eine Neugestaltung der getrennten Datenbanken mit Zeitreihenmetadaten und Zeitreihen zur Überwindung vorhandener Heterogenitäten nicht durchführbar, da diese Datenräume in ihrer jeweiligen Ausprägung die Grundlage des laufenden Forschungsbetriebes darstellten. Entsprechend mussten Wege gefunden werden, eine geeignete Lösung zu entwickeln und in den laufenden Forschungsbetrieb zu integrieren, *ohne* diesen zu unterbrechen oder zu beeinträchtigen.

### 9.3 Schwierigkeiten einer vollständigen Vorabdefinition

Die zu entwickelnde Lösung erforderte den Einsatz geeigneter Informationstechnologien und informationstechnologischer Konzepte zur effizienten Umsetzung der gegebenen Anforderungen. Eine vollständige Vorabdefinition, wie dies in geeigneter Weise erfolgen sollte, war dabei im gegebenen Kontext aufgrund der nachfolgend beschriebenen Faktoren erheblich erschwert.

#### 9.3.1 Einzubeziehende Kenntnisse und erforderlicher Zeitaufwand

Die für eine geeignete Lösungsdefinition einzubeziehenden Kenntnisse konnten weder von einzelnen Personen noch von einzelnen Wissenschaftsdisziplinen alleine bereitgestellt werden. Es war unabdingbar, sowohl das Anforderungswissen von Wissenschaftlern unterschiedlicher Disziplinen als Datennutzer wie das Fachwissen individueller Datenbereitsteller zu berücksichtigen, zu integrieren und in geeigneter Weise abzubilden. Voraussetzung war entsprechend die Zusammenführung der komplexen, fachspezifischen und auf eine Vielzahl unterschiedlicher Personen verteilten Kenntnisse um jeweils relevante Teilaspekte der Gesamtaufgabe. Eine vollständige und korrekte Vorabdefinition wurde dabei dadurch erschwert, dass für die einzubeziehenden Akteure eine Erfüllung ihrer jeweiligen wissenschaftlichen Aufgaben Vorrang besitzen musste. Eine umfassende theoretische Analyse von Anforderungen und Lösungsmöglichkeiten unter Einbeziehung aller relevanten Anforderungsbeitragenden war entsprechend aufgrund des hierfür erforderlichen Zeitaufwandes nicht mit den gegebenen Voraussetzungen vereinbar.

#### 9.3.2 Verschiedene fachspezifische Blickwinkel

Ferner war zu berücksichtigen, dass sich die erforderliche Kompetenz für die Beurteilung einzelner Ausschnitte der Gesamtaufgabe auf jeweils unterschiedliche Akteure verteilt. Letztlich können nur Datennutzer beurteilen, welche Funktionalitäten gemäss ihrer jeweiligen Bedürfnisse erforderlich und geeignet sind, während nur Datenbereitsteller über die jeweils erforderlichen Kenntnisse über die Organisation der von ihnen administrierten Datenräume - und damit für einen effizienten Zugang zu diesen - verfügen. Dabei konnte nicht davon ausgegangen werden, dass individuelle Datennutzer oder Datenbereitsteller umfassende Aussagen für *alle* Beteiligten ihrer jeweiligen Akteursgruppe treffen, sondern vielmehr jeweils nur Angaben gemäss ihres jeweiligen fachspezifischen Blickwinkels beitragen können. Mitglieder beider Akteursgruppen verfügen zudem in der Regel wiederum nur über limitierte Kenntnisse über alle einzubeziehenden Aspekte von Informationstechnologien und können entsprechend nur eingeschränkte Aussagen über die Umsetzung spezifischer Anforderungen treffen. Eine korrekte theoretische Beurteilung und Gewichtung aller für eine Lösung relevanten Aspekte war daher entsprechend erschwert.

#### 9.3.3 Unterschiedliches Hintergrundwissen und Fachvokabular

Schließlich darf nicht unterschätzt werden, dass jede Wissenschaft immer auch *esoterisch* ist und sich gegen andere Disziplinen durch eigene Konzepte und eigene Fachsprachen mit jeweils definierten Bedeutungen abgrenzt. Individuelle Datennutzer wie Datenbereitsteller verfügen über jeweils spezifisches, umfassendes Hintergrund- und Konzeptwissen sowie ein unterschiedliches Fachvokabular, so dass eine umfassende wechselseitige Verständigung zwischen den Anforderungsbeitragenden über die angestrebte Lösung erschwert wurde. Da sich zudem die beiden Gruppen der Datennutzer und Datenbereitsteller jeweils aus Wissenschaftlern einer Vielzahl unterschiedlicher Fachgebiete zusammensetzen, konnte überdies selbst *innerhalb* einer der beiden Akteursgruppen nicht von einer einfachen und erfolgversprechenden Kommunikation über die Anforderungen einer geeigneten Lösung ausgegangen werden.