

Verbundprojekt „Virtuelle Forschungsumgebung für die Sozioökonomische Berichterstattung“ (VFU soeb 3)

GEFÖRDERT VOM



Das Verbundprojekt „Virtuelle Forschungsumgebung für die Sozioökonomische Berichterstattung“ (VFU soeb 3) wurde vom Bundesministerium für Bildung und Forschung (BMBF) von Juni 2012 bis Juni 2014 mit rund 1 Mio. € gefördert. Das Projekt stützte sich auf konzeptionelle Vorarbeiten aus dem Teilprojekt „Kollaborative Datenauswertung und Virtuelle Arbeitsumgebung“ (Bartelheimer/ Schmidt 2011) im Rahmen des vom BMBF geförderten Projekts „Dritter Bericht zur sozioökonomischen Entwicklung Deutschlands – Durchführung einer Konzeptphase“.

Bei der Entwicklung der Virtuellen Forschungsumgebung kooperierte das Soziologische Forschungsinstitut Göttingen (SOFI) e.V. in einem interdisziplinären Projektverbund mit GESIS – Leibniz-Institut für Sozialwissenschaften e.V. (Mannheim), der Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG), dem Forschungsdatenzentrum der Bundesagentur für Arbeit (Nürnberg) im Institut für Arbeitsmarkt- und Berufsforschung (IAB) (FDZ-IAB), der Georg-August-Universität Göttingen Stiftung Öffentlichen Rechts (ausführende Stelle: Niedersächsische Staats- und Universitätsbibliothek (SUB)), den Forschungsdatenzentren der Rentenversicherung (FDZ-RV) und des Sozio-oekonomischen Panels (FDZ-SOEP), sowie der D-Grid gGmbH. Das SOFI als Forschungseinrichtung, die IT-

Dienstleister und die Dateneinrichtungen brachten ihre spezifischen Perspektiven in die Zusammenarbeit ein und waren dabei unterschiedlich eng an verschiedene Fachöffentlichkeiten angeschlossen.

Nach der Definition der Allianz-Initiative ist eine Virtuelle Forschungsumgebung „eine Arbeitsplattform, die eine kooperative Forschungstätigkeit durch mehrere Wissenschaftler an unterschiedlichen Orten zu gleicher Zeit ohne Einschränkungen ermöglicht. Inhaltlich unterstützt sie potenziell den gesamten Forschungsprozess – von der Erhebung, der Diskussion und weiteren Bearbeitung der Daten bis zur Publikation der Ergebnisse –, während sie technologisch vor allem auf Softwarediensten und Kommunikationsnetzwerken basiert.“¹ Die Projektpartner aus dem Bereich der Informationstechnologie und der Wissenschaftsdienstleistungen haben den Prototyp einer Virtuellen Forschungsumgebung entwickelt, der die kooperative Nutzung insbesondere sozial- und wirtschaftswissenschaftlicher Mikrodaten unterstützt, die von Forschungsdatenzentren bereitgestellt werden.

Nach Auffassung der Deutschen Forschungsgemeinschaft (DFG) und der Kommission Zukunft der Informationsinfrastruktur (KZII) können einsatzfähige VFU nur forschungsnah entwickelt werden (vgl. DFG 2010: 25). Die Architektur und die Funktionen einer VFU lassen sich nicht standardisieren. Sie sollen ausdrücklich die fachlich-inhaltliche Vielfalt unterstützen und „je nach fachspezifischen und individuellen Charakteristika strukturell sehr weit ausdifferenziert sein“ (KZII 2011: B74). Die Entwicklung der VFU *soeb 3* war daher eng an den Forschungsverbund „Dritter Bericht zur sozioökonomischen Entwicklung in Deutschland“ (*soeb 3*) angebunden, dem sie als Infrastruktur zur Verfügung gestellt und in dem sie projektbegleitend erprobt wurde. Die fachwissenschaftlichen Anforderungen an die VFU wurden in der Vorbereitung dieses Forschungsverbunds formuliert, und Wissenschaftler/innen aus dem Verbund dokumentierten ihre Erfahrungen im Rahmen von zwei Nutzungsstudien. Die Verknüpfung beider Projekte sollte einerseits eine notwendig forschungsnah Entwicklung des Infrastrukturprojekts ermöglichen und andererseits die wissenschaftliche Arbeit im Verbund wirksam unterstützen. Das Projekt sollte aber entsprechend der Zielsetzung der DFG (2013: 2 f.) zugleich ein praxistaugliches Resultat erzielen, das von einschlägigen Einrichtungen nachgenutzt werden kann. Die bisher festgestellte Beschränkung auf Lösungen für individuelle Projektziele sollte überwunden werden. Ziel war daher die Entwicklung eines VFU-Prototyps, der typische Workflows quantitativ-empirischer Sozial- und Wirtschaftswissenschaften unterstützt und eine ganze „Community“ potenzieller Nutzer/innen ansprechen kann.

¹ <http://www.allianzinitiative.de/de/handlungsfelder/virtuelle-forschungsumgebung.html>,
Stand: 07.10.2014.

In einer Entwicklungs- und Einführungsphase von 24 Monaten entstand der Prototyp eines VFU-Portals für die kollaborative Auswertung sozialwissenschaftlicher Mikrodaten, das mit ersten kollaborativen Werkzeugen in der Forschungspraxis erprobt wurde und mit Projektabschluss einsetzbar ist. Die Gesamtkoordination des Projekts lag beim SOFI und wurde von Dr. Peter Bartelheimer als Projektleiter und Sarah Cronjäger betreut. In Zusammenarbeit mit Dr. Tanja Schmidt wurden auf der sozialwissenschaftlichen Seite die fachwissenschaftlichen Anforderungen an eine VFU definiert und zwei Nutzungsstudien für die entwickelten Komponenten der VFU konzipiert und durchgeführt. Die Ergebnisse der zwei Nutzungsstudien wurden an die technischen Entwicklungen rückgekoppelt. Auf der informationstechnologischen moderieren Dr. Heike Neuroth und Timo Gnadt von der SUB unter Beteiligung von Prof. Dr. Uwe Schwiegelshohn von der D-Grid gGmbH eine technische Arbeitsgruppe, in der die Entwicklungsarbeit zwischen den beteiligten IT-Einrichtungen unter Einbeziehung der FDZ an einer gemeinsamen Architektur ausgerichtet wurde. Von der SUB wurde außerdem ein Konzept zur Langzeitarchivierung vorgelegt.

Unter Leitung von Prof. Dr. Ramin Yahyapour und Dr. Philipp Wieder haben Dr. Christian Boehme, Christof Pohl und weitere Mitarbeiter von der GWDG federführend die Entwicklung des VFU-Portals und des Hostings sowie die Integration der Middleware übernommen. Ein zentrales Portal für die VFU wurde mit Authentifizierung und rollenbasierter Autorisierung eingerichtet. Unter einer gemeinsamen Oberfläche integriert das VFU-Portal die Entwicklungen der GESIS (siehe unten) und weitere Werkzeuge zur Kollaboration. Hierzu gehören u.a. ein Forum, ein eigenes Wiki, eine Dateiverwaltung, ein gemeinsamer Projektkalender, ein Ankündigungs- und Absprachenportlet, ein Informationsportlet zur Planung von Teilnahmen an externen Tagungen, ein Publikationsportlet zur Dokumentation von Teilnahmen an externen Tagungen und geplanten und realisierten Veröffentlichungen, sowie eine Archivfunktion zur Speicherung und zum Durchsuchen von Datensätzen und Metadaten. Die Datenverwaltung unterstützt schwerpunktmäßig die Archivierung, Dokumentation und Nachnutzung bzw. kollaborative Nutzung von Auswertungssyntax. Neben Syntaxdateien und (freigegebenen) Forschungsdaten werden über sie auch alle anderen Dateiformate, etwa Outputdateien, Tabellen- oder Textdokumente, formatunabhängig abgelegt und mittels Inhaltssuche wie über Metadaten auffindbar verwaltet. Im Rahmen der bestehenden Nutzungsberechtigungen können logische Verknüpfungen zwischen den verschiedenen Datenarten (z.B. zwischen Syntax und Forschungsdaten) hergestellt werden.

Eine entscheidende Voraussetzung für die Nutzung einer gemeinsamen Arbeitsumgebung besteht darin, dass die Daten anhand von Metadaten gut beschrieben werden. Bei der GESIS wurden unter der Leitung von Uwe Jensen und der Mitarbeit von Stefan Schweers die Metadatenelemente definiert, die in der VFU benötigt werden. Die Beziehungen der Metada-

tenelemente, die von den vertretenen FDZ verwendet und eingebracht wurden, zu bestehenden Metadatenstandards wurden geprüft und festgelegt. Für die *VFU soeb 3* wurde ein eigenes Metadatenschema spezifiziert, das sich an den Standard der Data Documentation Initiative (DDI), der sich für Forschungsdaten international etabliert hat, orientiert. Neben der Entwicklung eines Metadatenschemas wurden bei der GESIS unter Leitung von Dr. Brigitte Mathiak und der Mitarbeit von Zeljko Carevic die kollaborativen IT-Werkzeuge, insbesondere die Syntax- und Metadateneditoren, entwickelt.

Die beteiligten Forschungsdatenzentren, wie das FDZ der BA im IAB (Stefan Bender, David Schiller), das FDZ-SOEP (Dr. Jan Goebel) und das FDZ-RV (Dr. Michael Stegmann, Tatjana Mika) haben ihre Daten- und Metadatenkompetenzen, ihr Wissen aus den Bereichen Langzeitarchivierung, die organisatorischen und technischen Voraussetzungen des Zugangs zu ihren Mikrodaten (Remote Access und/oder Job Submission) sowie ihre datenschutzrechtliche Anforderungen für den Umgang mit Mikrodaten und den Ablauf wissenschaftlichen Arbeitens mit sensiblen Mikrodaten in die Entwicklung der VFU eingebracht. Sie vermittelten zwischen fachwissenschaftlichen Anforderungen und Wissenstechnologien.

Das Verbundprojekt *VFU soeb 3* sollte die Erfolgsaussichten für neue Zugangswege zu Mikrodaten in Dateneinrichtungen verbessern und Lösungen für deren Unterstützung in einer VFU skizzieren. Entwicklungsaufgaben für eine insbesondere um Werkzeuge für Remote Access erweiterte Version der VFU wurden spezifiziert und an einem Proof of Concept demonstriert.

Die Abschlussveranstaltung des Projekts *VFU soeb 3* fand am 24. Januar 2014 in Berlin statt. Der Projektverbund präsentierte die erste operativ genutzte Version der VFU einer breiteren Fachöffentlichkeit und stellte sie zur Diskussion. Die Tagung wurde auf der Projektwebseite dokumentiert.²

Dass die *VFU soeb 3* als operativ einsetzbarer Prototyp entwickelt wurde und bei den Pilotanwendern des Forschungsverbunds *soeb 3* hohe Akzeptanz findet, zeigt den wissenschaftlichen Erfolg des Projekts. Damit hat das Projekt zumindest deutschlandweit eine Pionierrolle bei der Erstellung von virtuellen Forschungsumgebungen für die quantitativ-empirisch forschenden Sozial- und Wirtschaftswissenschaften und kann bei der Weiterentwicklung und Nutzung von VFUen in diesem Bereich durch die erbrachten Entwicklungen und Erfahrungen unterstützend wirken.

Die entwickelte Gesamtarchitektur der VFU erlaubt eine einfache konzeptuelle Anpassung für andere Fach-Communities oder Forschungsschwerpunkte und deren Workflows. Die entwickelte VFU soll daher künftig anderen Anwender/inne/n in einer erweiterten Testphase

² <http://www.soeb.de/vfu-soeb-3/abschlusstagung/>, Zugriff am 07.10.2014.

zur Verfügung gestellt werden. Dafür wird durch Eigenmittel der Projektpartner im Anschluss an die Projektförderung ein entsprechendes Betriebsmodell und eine Demoversion entwickelt.

Weitere Informationen finden Sie unter: <http://www.soeb.de/vfu-soeb-3/>

Kontakt

Dr. Peter Bartelheimer, Soziologisches Forschungsinstitut Göttingen (SOFI) an der Georg-August-Universität, Friedländer Weg 31, 7085 Göttingen, Tel. +49 551 52205-51; Fax +49 551 5220588, E-Mail: peter.bartelheimer@sofi.uni-goettingen.de

Sarah Cronjäger, M.A., Soziologisches Forschungsinstitut Göttingen (SOFI) an der Georg-August-Universität, Friedländer Weg 31, 37085 Göttingen, Tel.: +49 551 52205-38, E-Mail: Sarah.Cronjaeger@sofi.uni-goettingen.de

Literatur

Bartelheimer, Peter/ Schmidt, Tanja (2011): Modellprojekt „Kollaborative Datenauswertung und Virtuelle Arbeitsumgebung“ – VirtAug. Abschlussbericht. Download unter: http://www.soeb.de/fileadmin/redaktion/downloads/VirtAug/VirtAug_Abschlussbericht.pdf, Stand: 10.10.2014.

Deutsche Forschungsgemeinschaft (DFG) (2010): Informationsverarbeitung an Hochschulen – Organisation, Dienste und Systeme. Empfehlungen der Kommission für IT-Infrastruktur 2011-2015. Bonn. Download: http://www.hof.uni-halle.de/dateien/pdf/DFG_Empfehlungen_IT_2011_2015.pdf, Stand: 08.10.2014.

Deutsche Forschungsgemeinschaft (DFG) (2013): Vernetzung – Austausch – Nachnutzung. Workshop zur Präsentation von Projekten aus dem Förderbereich Informationsmanagement der DFG. Bonn.

Kommission Zukunft der Informationsinfrastruktur (zitiert als KZII) (2011): Gesamtkonzept für die Informationsinfrastruktur in Deutschland. Empfehlungen der Kommission Zukunft der Informationsinfrastruktur im Auftrag der Gemeinsamen Wissenschaftskonferenz des Bundes und der Länder. April 2011. Download: http://www.leibniz-gemeinschaft.de/fileadmin/user_upload/downloads/Infrastruktur/KII_Gesamtkonzept.pdf, Stand: 07.10.2014.