

Abschließender Sachbericht

**Titel des Vorhabens:
SowiDataNet – Forschungsdatenverbund
für Sozial- und Wirtschaftswissenschaften:
Infrastruktur für die eigenständige
Archivierung und Distribution
quantitativer Forschungsdaten der
Sozial- und Wirtschaftswissenschaften**

Leibniz-Einrichtung: GESIS – Leibniz-Institut für Sozialwissenschaften
Aktenzeichen: SAW-2014-GESIS-2
Projektlaufzeit: 01.04.2014 – 31.12.2017
Ansprechpartner: Wolfgang Zenk-Möltgen

Inhaltverzeichnis

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Executive summary _____ | 3 |
| Ausgangsfragen und Zielsetzung des Vorhabens _____ | 4 |
| Entwicklung der durchgeführten Arbeiten einschließlich Abweichungen vom ursprünglichen Konzept, wissenschaftliche Fehlschläge, Probleme in der Vorhabenorganisation oder technischen Durchführung _____ | 6 |
| Darstellung der erreichten Ergebnisse und Diskussion im Hinblick auf den relevanten Forschungsstand, mögliche Anwendungsperspektiven und denkbare Folgevorhaben _____ | 7 |
| Stellungnahme, ob Ergebnisse des Vorhabens wirtschaftlich verwertbar sind und ob eine solche Verwertung erfolgt oder zu erwarten ist; Angaben zu möglichen Patenten oder Industriekooperationen _____ | 8 |
| Angabe der Beiträge von möglichen Kooperationspartnern im In- und Ausland, die zu den Ergebnissen des Vorhabens beigetragen haben _____ | 9 |
| Qualifikationsarbeiten, die im Zusammenhang mit dem Vorhaben entstanden sind _____ | 9 |
| Institutionen und Personen, die am Projektvorhaben mitgewirkt haben _____ | 9 |
| GESIS – Leibniz-Institut für Sozialwissenschaften _____ | 9 |
| WZB – Wissenschaftszentrum Berlin für Sozialforschung _____ | 9 |
| DIW Berlin – Deutsches Institut für Wirtschaftsforschung e.V. _____ | 9 |
| ZBW – Leibniz-Informationszentrum Wirtschaft _____ | 10 |
| Liste der Publikationen aus dem Vorhaben _____ | 10 |
| Aufsätze und Projektberichte _____ | 10 |
| Vorträge _____ | 10 |
| Darstellung der Maßnahmen zur Sicherung und Verfügbarmachung der im Vorhaben produzierten Forschungsdaten _____ | 11 |
| Liste möglicher Pressemitteilungen und Medienberichte _____ | 12 |
| Anhang: Überblick über die Arbeitsgruppen _____ | 12 |
| AP1 Anforderungsanalyse (Datenveröffentlichung/-kuratierung/-suche) Leitung: WZB; Partner: GESIS, DIW, ZBW _____ | 12 |
| AP2 Organisationsmodell (Policies/Nutzungsbedingungen/AGBs/Institutsverträge) Leitung: ZBW; Partner: DIW, WZB _____ | 13 |
| AP3 Metadaten-Strukturen Leitung: GESIS; Partner: WZB _____ | 13 |
| AP4 Softwareentwicklung (Repositoriumssoftware/Institutionelle Ansicht) Leitung: GESIS; Partner: ZBW _____ | 14 |
| AP5 Prototypentest/Beispielinhalte Leitung: WZB; Partner: DIW _____ | 14 |
| AP6 Transfer/Soziale Medien/Kommunikation Leitung: DIW; Partner: GESIS, WZB, ZBW _____ | 14 |

Executive summary

Basis für wissenschaftliche Erkenntnisse und ihre Veröffentlichung sind Forschungsdaten. Diese müssen für Replikationszwecke und Nachnutzung für neue Forschungsfragen frei und langfristig verfügbar sein, damit eine Transparenz im Forschungsprozess sichergestellt ist. Ziel des Vorhabens SowiDataNet ist daher eine gemeinsame Infrastruktur für eine nutzerfreundliche Integration dezentral erzeugter Forschungsdaten der Sozial- und Wirtschaftswissenschaften in Deutschland. Dies ermöglicht teilnehmenden Institutionen die Verwendung einheitlicher Standards für die Archivierung und Dokumentation, macht die Forschungsdaten zitierbar und zugänglich, und ermöglicht innovative Recherchemöglichkeiten. SowiDataNet hat dabei besonders Forschungseinrichtungen ohne eigene Forschungsdateninfrastruktur im Fokus.

Durch den Zusammenschluss der vier Partner GESIS, DIW Berlin, WZB und ZBW besteht der Verbund aus je zwei Einrichtungen mit überwiegendem Forschungsschwerpunkt und zwei mit überwiegendem Infrastrukturschwerpunkt. Hierdurch wird die Entwicklung der Infrastruktur an den Bedarfen und methodologischen Standards der Sozial- und Wirtschaftswissenschaften orientiert.

Im Projekt wurden durch eine nutzerorientierte Bedarfsermittlung, eine Literaturanalyse, sowie durch eine Reihe qualitativer Experteninterviews die Anforderungen an eine Forschungsdaten-Infrastruktur und ihre Funktionen ermittelt. Das Repository SowiDataNet wurde mithilfe der freien Software DSpace für Erfassung, Speicherung und Distribution von digitalen Ressourcen realisiert. Dieses konnte auf Vorarbeiten von GESIS aufsetzen, wo mit dem Repository ein Forschungsdatenrepository existiert, in dem Einzelforscherinnen ihre Daten eigenständig dokumentieren können. In SowiDataNet wurden die zusätzlichen Funktionalitäten nach Maßgabe der Anforderungsanalyse erarbeitet und implementiert, sowie anhand verschiedener exemplarischer Forschungsdatensätze aus der Praxis der Partnerinstitute getestet. Dabei wurden besonders Forschungsdatensätze mit verschiedenen Formaten berücksichtigt und komplexe Anforderungen an Metadaten und Usability der Benutzeroberfläche evaluiert. Im Projekt wurden regelmäßige virtuelle Meetings, verschiedene Arbeitsgruppen und Veranstaltungen durchgeführt, so dass die Tests nicht nur durch die Kooperationspartner, sondern auch mit projektexternen Teilnehmenden stattfanden.

Das im Projekt entwickelte Repository SowiDataNet erlaubt es nun, dass wissenschaftliche Einrichtungen ihre Daten, Metadaten und relevanten Dokumente eigenständig in das Repository einpflegen, sie systematisch archivieren und zeitnah der Wissenschaftsgemeinschaft zur Verfügung stellen können. Die Implementierung der „institutionellen Ansicht“ hat für teilnehmende Institutionen den Vorteil, dass die von ihnen verwalteten Forschungsdaten auch in ihrem eigentlichen Webangebot unter Verwendung ihres eigenen Corporate Designs dargestellt werden können. Institutionen haben hiermit eine Möglichkeit zur Einbindung von SowiDataNet in die Prozesse des eigenen institutionellen Forschungsdatenmanagements. Dabei wird eine persistente Identifizierung von Forschungsdaten über DOIs geleistet. Das Metadatenschema erlaubt eine Dokumentation auch komplexer Forschungsdatenobjekte, wobei verschiedene Stufen von Granularität berücksichtigt werden und eine Kompatibilität zu wichtigen Metadatenstandards wie Dublin Core, DataCite und DDI sowie eine Verwendung einschlägiger Thesauri, Klassifikationen und kontrollierter Vokabulare gewährleistet ist. Des Weiteren ist ein Versionierungskonzept für das Management von Updates der Forschungsdaten in SowiDataNet implementiert.

GESIS hat im Rahmen umfangreicher Tests entschieden, das entwickelte SowiDataNet-Repository in seine Services zu integrieren. Dazu wird nach der gegenwärtigen Beta-Phase der volle Umfang der Dienstleistung im Laufe des Jahres 2018 zur Verfügung stehen. In einer Reihe von Publikationen und Konferenzbeiträgen wurde SowiDataNet in der Community bekannt gemacht. Dies mündete in Interessensbekundungen verschiedenster Einrichtungen aus den Sozial- und Wirtschaftswissenschaften.

Ausgangsfragen und Zielsetzung des Vorhabens

Forschungsdaten sind in vielen Disziplinen die Grundlage für neue wissenschaftliche Erkenntnisse und bilden die Basis der daraus resultierenden Veröffentlichungen. Gerade in Zeiten einer immer stärkeren Digitalisierung, in der die überwiegende Anzahl von Forschungsdaten in elektronischer Form vorliegt, kommt der flexiblen Datendistribution und der Nachnutzung von Forschungsdaten eine herausragende Bedeutung zu. Ihre freie und langfristige Verfügbarkeit sichert den Erkenntnisgewinn und stellt die notwendige Transparenz und Nachvollziehbarkeit im Forschungsprozess sicher¹. Die Möglichkeit zur Replikation trägt zur Vermeidung redundanter Untersuchungen bei, sorgt für eine fachliche Qualitätskontrolle und ist darüber hinaus eine nützliche Basis für die Ausbildung in der Wissenschaft². Ferner trägt die Verfügbarkeit von Forschungsdaten auch dazu bei, dass die Primärforscherinnen öfter zitiert werden³, womit ein wesentlicher Anreiz für die teils durchaus zeitaufwändige Dokumentation der Daten gegeben wird.

Für eine Umsetzung dieser Empfehlungen von diversen Förder- und Wissenschaftsorganisationen existierten jedoch bislang für die Sozial- und Wirtschaftswissenschaften kaum adäquate Infrastrukturen. An dieser Stelle setzt der Projektverbund SowiDataNet an.

Eines der übergeordneten Projektziele ist die Integration dezentral erzeugter Forschungsdaten unter Nutzung einer gemeinsamen Infrastruktur. Dies hat den Vorteil, dass Forschungsdaten der Sozial- und Wirtschaftswissenschaften in Deutschland nach einheitlichen Standards archiviert und dokumentiert werden, zitierbar und zugänglich gemacht werden können. Die Zusammenführung von umfassend dokumentierten Forschungsdaten aus unterschiedlichen Quellen verbessert zudem die Auffindbarkeit der Daten und ermöglicht innovative Recherchemöglichkeiten, bspw. nach thematischen Schlagworten oder nach unterschiedlichen Erhebungsmethoden. Mit dem Projekt werden darüber hinaus die Möglichkeiten einer effektiven Langzeitarchivierung der Forschungsdaten aus den Sozial- und Wirtschaftswissenschaften substantiell verbessert (durch eine perspektivische Anbindung an das Datenarchiv bei GESIS). Da der Kontakt zu den Primärforschern nur selten dauerhaft sicher gestellt werden kann, ist es schließlich Ziel des Projekts, Forschungsdaten über SowiDataNet in einer Form zu dokumentieren, die es erlaubt, sie später auch ohne Unterstützung durch die Datenproduzentinnen nachzuvollziehen und nachnutzen zu können.

Insgesamt zielt das Projekt somit auf den Aufbau und die Etablierung eines digitalen Forschungsdaten-Repositorys für die Sozial- und Wirtschaftswissenschaften, welches Wissenschaftlern eine sichere Archivierung und Weitergabe ihrer Forschungsdaten auf nutzungsfreundliche Art und Weise ermöglicht. Dabei richtet sich das Angebot besonders an Einrichtungen, die kein eigenes separates Repository aufbauen und betreiben wollen oder können. Solche Einrichtungen können das Angebot für ihre wissenschaftlichen Mitarbeitenden nutzen, um diese bei der Durchführungen eines Forschungsdatenmanagements nach dem Stand der Wissenschaft zu unterstützen.

Um ein möglichst großes Spektrum an Daten aufzunehmen, wird der Begriff „Forschungsdaten“ durch das Projekt weit gefasst. So können sowohl Primärdaten in das Repository aufgenommen werden, die aufgrund der Nutzung empirischer Forschungsinstrumente neu produziert werden, als auch Sekundärdaten, die z.B. aus vorhandenen Quellen in neue Datenbanken transformiert werden. Darüber hinaus können ebenfalls Kombinationen von Primär-

¹ Vgl. Jasny, Barbara R.; Chin, Gilbert; Chong, Lisa; Vignieri, Sacha: Again, and Again, and Again ... In: Science 6060 (2011), S. 1225.

² Vgl. King, Gary: Replication, Replication. In: Political Science and Politics 3 (1995), S. 443– 499.

³ Vgl. Piwowar, Heather A.; Day, Roger S., Fridsma, Douglas B.: Sharing Detailed Research Data Is Associated with Increased Citation Rate. In: PLoS ONE 3 (2007), e308.

und Sekundärdaten sowie die für Datentransformationen und -analysen benutzten Routinen oder Skripte berücksichtigt werden. Ergänzend können Dokumente bereitgestellt werden, welche die Nachnutzung der Daten erleichtern (etwa Fragebögen, Codebooks oder Technical Reports).

Im Fokus von SowiDataNet liegen zunächst Projekte mit fehlenden oder geringen personellen und finanziellen Ressourcen für die Dokumentation und Archivierung von Forschungsdaten. Dies sind in der Regel mittlere und kleine Forschungs- oder Promotionsprojekte. Die fachliche Zielgruppe ist die sozial- und wirtschaftswissenschaftliche Forschungsgemeinschaft in Deutschland. Mit den adressierten Disziplinen werden Fächer angesprochen, die zusammen genommen die meisten Studierendenzahlen und dazu entsprechende Lehrstuhl-Strukturen in Deutschland auf sich vereinen.

Bei der Zielgruppe ist zudem zwischen Datengeberinnen und Datennutzern zu differenzieren. Datengeberinnen sollen über SowiDataNet ihre Forschungsdaten in einfacher Weise dokumentieren, archivieren und anderen zugänglich machen können. Datennutzer sollen mittels der Metadaten recherchieren, Daten für Reanalysen oder neue Forschungsvorhaben nutzen können und mit SowiDataNet einen zentralen Zugang für die Arbeit mit sozial- und wirtschaftswissenschaftlichen Forschungsdaten erhalten.

Um eine flexible Datendistribution und die wissenschaftliche Nachnutzung von Forschungsdaten in den Sozial- und Wirtschaftswissenschaften zu fördern, startete das Datenarchiv von GESIS bereits Anfang 2012 die Entwicklung von datorium (<https://datorium.gesis.org/xmlui/>) - einem digitalen Forschungsdatenrepositorium, in dem Einzelforscherinnen ihre Daten eigenständig dokumentieren können, um sie im Anschluss von GESIS veröffentlichen zu lassen⁴. Seit Januar 2014 ist das Repositorium online und erfährt seitdem eine zunehmende Nutzung. Im Laufe der Betriebsphase von datorium wurde jedoch schnell deutlich, dass zusätzlich eine digitale Arbeitsumgebung benötigt wird, die speziell auf die Bedürfnisse eines institutionellen Forschungsumfeldes zugeschnitten ist. Aus diesem Grunde schloss sich der Projektverbund SowiDataNet zusammen, um die spezifischen Anforderungen institutioneller Nutzer hinsichtlich eines Datenrepositoriums zu ermitteln und umzusetzen. So wird SowiDataNet mit besonderer Konzentration auf Forschungseinrichtungen ohne eigene Forschungsdateninfrastruktur weiterentwickelt, während datorium vorerst weiterhin auf Einzelforscherinnen fokussiert. Durch diese Vorgehensweise können die Bedarfe aller Nutzergruppen optimal erfasst und berücksichtigt werden. Nach Abschluss der Projektphase und einem produktiven Einsatz von SowiDataNet wird datorium und der darin enthaltene Datenbestand in SowiDataNet integriert werden. Dadurch entstehen ein Service und ein System, welches für verschiedene Nutzergruppen zugänglich ist und langfristig eine übergreifende und nutzerfreundliche Recherche ermöglicht. Dies steuert einer weiteren Zerklüftung der Datenlandschaft in Deutschland wirksam entgegen.

⁴ Vgl. Linne, Monika: Sustainable data preservation using datorium: facilitating the scientific ideal of data sharing in the social sciences. In: *Proceedings of the 10th International Conference on Preservation of Digital Objects*, edited by José Borbinha, Michael Nelson, and Steve Knight. Lisbon: 2013, S. 150-155.

Entwicklung der durchgeführten Arbeiten einschließlich Abweichungen vom ursprünglichen Konzept, wissenschaftliche Fehlschläge, Probleme in der Vorhabenorganisation oder technischen Durchführung

Den Startpunkt des Projektes im Juni 2014 bildete eine nutzerorientierte Bedarfsermittlung, in deren Rahmen ein Auftaktworkshop am Deutschen Institut für Wirtschaftsforschung (DIW), eine Literaturanalyse sowie eine Reihe qualitativer Experteninterviews durchgeführt wurden. Die Ergebnisse und die daraus abgeleiteten Anforderungen ergaben, dass das von GESIS betriebene Repositorium datorium bereits grundlegende technische Anforderungen von SowiDataNet erfüllt. Entsprechend wurde auch SowiDataNet mit DSpace aufgesetzt, einer freien Software für die Erfassung, Speicherung und Distribution von digitalen Ressourcen. Daher war das Entwicklungsteam in der vorteilhaften Lage, sehr zügig einen ersten Prototyp von SowiDataNet als eine technische und inhaltliche Weiterentwicklung von datorium zu implementieren. Dies hatte den Vorteil, dass doppelte Entwicklungen vermieden und Synergien effizient genutzt werden konnten.

Der erste Prototyp wurde Anfang 2015 fertiggestellt und anhand verschiedener exemplarischer Forschungsdatensätze aus der Praxis der Partnerinstitute getestet. Geprüft wurde die Bedienbarkeit und Eignung für Datensätze unterschiedlichen Formats sowie die Usability der Benutzeroberfläche. Die Erkenntnisse daraus wurden im Detail dokumentiert und flossen in die Weiterentwicklung des Prototyps ein.

Während des Projektverlaufs wurden regelmäßig virtuelle Meetings in verschiedenen Arbeitsgruppen durchgeführt, um den Projektfortschritt und die Zusammenarbeit der Partner zu steuern. Ein zusammenfassender Überblick mit Beschreibung der Arbeitsgruppen und ihren zentralen Ergebnissen findet sich im Anhang. Zusätzlich wurden, wie im Projektantrag vorgesehen, folgende Veranstaltungen durchgeführt:

| <i>Datum</i> | <i>Partner</i> | <i>Meeting</i> |
|----------------|----------------|--------------------------------------------------------------------------------------------------|
| 22.04.2014 | WZB | Kickoff-Meeting |
| 23.06.2014 | DIW | Projektworkshop |
| 24.10.2014 | GESIS | Interimsworkshop |
| 13.05.2015 | ZBW | Interimsworkshop |
| 02.10.2015 | DIW | Interimsworkshop |
| 28./29.01.2016 | GESIS | Interimsworkshop |
| 23./24.06.2016 | ZBW | Interimsworkshop |
| 14./15.11.2016 | WZB | Interimsworkshop |
| 10.02.2017 | DIW | Präsentation und Live-Demo mit 65 projektexternen Teilnehmenden aus interessierten Einrichtungen |

2017 wurde die Testphase auf institutionelle externe Beta-Nutzerinnen erweitert. In dieser Phase war die Funktionalität des Repositoriums bereits gegeben. Die Rückmeldungen aus dieser Testphase waren ausschließlich positiv.

Durch noch zur Verfügung stehende Mittel konnte eine kostenneutrale Laufzeitverlängerung des Vorhabens bis Ende Dezember 2017 erfolgen. Diese wurde für weitere Arbeiten in der Projektleitung verwendet, sowie für eine Konsolidierung der technischen Infrastruktur, eine

Beratung potentieller institutioneller Kunden und eine Vorbereitung für die Fusion des Sowi-DataNet mit dem bereits bestehenden GESIS datorium-Angebot.

Darstellung der erreichten Ergebnisse und Diskussion im Hinblick auf den relevanten Forschungsstand, mögliche Anwendungsperspektiven und denkbare Folgevorhaben

In SowiDataNet erhalten wissenschaftliche Einrichtungen die Möglichkeit, ihre Daten, Metadaten und relevanten Dokumente eigenständig in das Repositorium einzupflegen sowie ihre Daten mit dazugehörigen Publikationen zu verlinken. Darüber hinaus können sie anhand von Zugangsklassen selbst definieren, unter welchen Bedingungen – z.B. freie Verfügbarkeit bis hin zu Embargofristen – sie den Zugriff auf ihre Daten erlauben. Durch dieses Vorgehen obliegt die Kontrolle über die Datenweitergabe den Forscherinnen. Grundsätzlich ist es jedoch stets das Ziel des Repositoriums, die Daten verfügbar zu machen, weshalb eine Embargofrist auf maximal zwei Jahre festgelegt wurde.

Den Instituten erlaubt SowiDataNet ihre Forschungsergebnisse systematisch zu archivieren und zeitnah der Wissenschaftsgemeinschaft zur Verfügung zu stellen. Im Projekt wurde auch die Implementierung der „institutionellen Ansicht“ umgesetzt, die es teilnehmenden Institutionen ermöglicht, die von ihnen verwalteten Forschungsdaten auch in ihrem eigentlichen Webangebot darzustellen. Dabei ist es möglich, das eigene Corporate Design zu verwenden, und dynamisch, d.h. jederzeit aktuell die Inhalte aus dem SowiDataNet-Repositorium im eigenen Webangebot zu verwenden. SowiDataNet bietet zudem die vordefinierte Rolle eines institutionellen Kurators, mittels der sich das System in die Workflows des institutionellen Forschungsdatenmanagements integrieren lässt.

Zur sicheren und persistenten Identifizierung von Forschungsdaten nutzt SowiDataNet den Service der Registrierungsagentur für Sozial- und Wirtschaftsdaten da|ra zur DOI-Vergabe (<http://www.da-ra.de/>). Der DOI-Name identifiziert ein Forschungsprojekt in SowiDataNet unmittelbar und stellt die persistente Verknüpfung mit strukturierten Metadaten sicher. Dies ist eine zentrale Voraussetzung für eine dauerhafte Verbindung von Forschungsergebnissen in Form von Publikationen und den ihnen zugrunde liegenden Primärdaten.

Das Metadatenschema von SowiDataNet enthält obligatorische Kernelemente, die zur Beschreibung eines Datensatzes vorhanden sein müssen: Titel, Primärforscher, Institution, Jahr der Veröffentlichung und Verfügbarkeit. Darüber hinaus können optionale Metadatenelemente angegeben werden, um die Daten präziser zu erfassen, z.B. das Fachgebiet, das verwendete Erhebungsverfahren oder die Anzahl der Variablen. Da die Sozial- und Wirtschaftswissenschaften ein disziplinäres Umfeld sind, in dem die Forschungsdesigns sich häufig durch Komplexität kennzeichnen (z.B. Mixed-Methods-Ansätze), bietet SowiDataNet weiterhin die Option, bei der Metadatenbeschreibung verschiedene Granularitätsstufen zu berücksichtigen: Bei der Dokumentation der Forschungsdaten unterscheidet das Repositorium zwischen einer Projekt- und Objektebene. Auf der Projektebene können Angaben zum Forschungsdatenprojekt als Ganzes gemacht werden. Sollten sich einzelne Datenobjekte von dieser allgemeinen Projektebene – z.B. durch abweichende Methodik – unterscheiden, ist es auch möglich, objektspezifische Metadaten zu vergeben.

Für sämtliche Metadatenelemente werden die jeweiligen Namen, Definitionen, Attribute, ggf. Bedingungen, Kardinalitäten (Anzahl), sowie Wertebereiche definiert, stellenweise wird auf weitere anzuwendende Standards verwiesen. Das Metadatenschema von SowiDataNet ist kompatibel zum DDI Codebook-Standard der Data Documentation Initiative (<http://www.ddialliance.org/>), sowie zu den Metadatenschemata der Datenregistrierungsagentur da|ra (<http://www.da-ra.de/home/>) und DataCite (<https://www.datacite.org/>). Soweit

möglich, werden Elemente der Dublin Core Metadata Initiative (<http://dublincore.org/>) verwendet.

Darüber hinaus wurde ein Versionierungskonzept für Forschungsprojekte umgesetzt, welches insbesondere für Forschungsdaten von institutionellen Datengeberinnen von Bedeutung ist, da diese in der Regel die Datensätze längerfristig pflegen und Fehlerkorrekturen vornehmen. Hierbei ist zu beachten, dass im System auf die jeweiligen Vorgänger- bzw. Nachfolgeversionen der Daten referenziert wird, damit stets zur gewünschten Version navigiert werden kann und Irritationen bei nachnutzenden Wissenschaftlern vermieden werden.

Eine spezielle Herausforderung lag schließlich im Bereich der Anwendung von Thesauri und Klassifikationen, welche sowohl für die Sozial- als auch Wirtschaftswissenschaften geeignet sein müssen. Deren Eignung ist besonders wichtig, um adäquate Funktionen zur Strukturierung und inhaltlichen Erschließung von Forschungsprojekten anzubieten. Insbesondere für die Nachnutzung und Auffindbarkeit von Forschungsdaten ist dies von zentraler Bedeutung. Das Repositorium nutzt daher für das Metadatenschema eine Reihe von Thesauri (Standard Thesaurus Wirtschaft STW; Thesaurus Sozialwissenschaft TheSoz) und kontrollierte Vokabulare, siehe dazu das Kapitel „Verwendung von kontrollierten Vokabularen in SowiDataNet“ aus dem SowiDataNet Metadatenschema (Jensen/Linne 2017).

GESIS hat im Rahmen der Beta-Tests bereits entschieden, das entwickelte SowiDataNet-Repositorium in seine Services zu integrieren. Dazu ist gegenwärtig eine Beta-Phase gestartet, die es interessierten Institutionen erlaubt, den vollen Umfang der Dienstleistung zu testen. Im Laufe des Jahres 2018 wird dieser Service in einen Echtbetrieb übergehen und kann dann von Institutionen nach Abschluss eines Servicevertrags mit GESIS zur Publikation von Forschungsdaten genutzt werden. Dabei wird zurzeit an einer Fusion mit dem bereits länger verfügbaren Dienst datorium gearbeitet. Weitere Informationen finden sich unter <http://sowidatanet.de/xmlui/>.

Die Nutzung von SowiDataNet ist zurzeit ausschließlich auf institutionelle Einrichtungen beschränkt, die ihre Forschungsdaten zu Nachnutzungszwecken publizieren. In der Projektphase von SowiDataNet wurde allerdings darüber hinaus der Bedarf eines institutsinternen Forschungsdatenmanagements identifiziert. Hieraus ergibt sich eine mögliche Weiterentwicklung, die in einem Folgeprojekt bearbeitet werden könnte: Existierende institutionelle Forschungsdaten, die (vorerst) nicht zur Veröffentlichung bestimmt sind, sowie Daten aus laufenden Projekten sollten institutsintern verwaltet werden können. Auch solche Daten müssen praxisnah gemanagt werden, ohne dabei Datenverluste zu riskieren oder auf Dauer Kontextinformationen zu verlieren, welche z. B. für Replikationszwecke oder Anschlussstudien benötigt werden. Diesbezüglich wäre zu erörtern, welche Managementprozesse und -aufgaben unterstützt werden müssen und wie darüber hinaus eine Langfristsicherung institutsinterner Daten unter Berücksichtigung datenschutzrelevanter Aspekte umsetzbar wäre.

Stellungnahme, ob Ergebnisse des Vorhabens wirtschaftlich verwertbar sind und ob eine solche Verwertung erfolgt oder zu erwarten ist; Angaben zu möglichen Patenten oder Industriekooperationen

Die Ergebnisse des Vorhabens ermöglichen den Aufbau der geplanten Infrastruktur. GESIS bietet die Services innerhalb der Infrastruktur zurzeit in einer Beta-Phase an und plant einen produktiven Betrieb. Eine wirtschaftliche Verwertung ist aus Sicht der Projektpartner nicht möglich und ist in absehbarer Zukunft auch nicht zu erwarten. Die im Rahmen der Infrastruktur angebotenen Services werden jedoch kostenpflichtig angeboten, um so einen Beitrag zur

Deckung der durch Angebot und Nutzung des Services entstehenden Kosten zu erzielen. Die aktuelle SowiDataNet-Preisliste findet sich unter http://sowidatanet.de/xmlui/static/files/SowiDataNet_Preise_2018-06-14_deutsch.pdf.

Es sind keine Patente oder Industriekooperationen entstanden.

Angabe der Beiträge von möglichen Kooperationspartnern im In- und Ausland, die zu den Ergebnissen des Vorhabens beigetragen haben

Wir danken allen Personen, die an den Nutzertests und Workshops teilgenommen haben und damit zur Erreichung der Projektergebnisse beigetragen haben, deren Namen wir aus Datenschutzgründen aber hier nicht nennen können.

Qualifikationsarbeiten, die im Zusammenhang mit dem Vorhaben entstanden sind

Zum Vorhaben sind keine Qualifikationsarbeiten entstanden.

Institutionen und Personen, die am Projektvorhaben mitgewirkt haben

GESIS – Leibniz-Institut für Sozialwissenschaften

Dülmer, Hermann
Haddou ou Moussa, Karima
Jensen, Uwe
Klas, Claus-Peter
Linne, Monika
Rudolph, Julia
Wasner, Catharina
Wira-Alam, Andias
Zenk-Möltgen, Wolfgang

WZB – Wissenschaftszentrum Berlin für Sozialforschung

Droß, Patrick
Habich, Roland
Kubaty, Paul
Naujoks, Julian
Nix, Sebastian

DIW Berlin – Deutsches Institut für Wirtschaftsforschung e.V.

Fecher, Benedikt
Fräßdorf, Mathis
Löwe, Peter
Sieggers, Rainer
Wagner, Gert G.

ZBW – Leibniz-Informationszentrum Wirtschaft

Borst, Timo
Neubert, Joachim
Ott, Konstantin
Toepfer, Ralf
Vlaeminck, Sven

Liste der Publikationen aus dem Vorhaben

Aufsätze und Projektberichte

Fecher, B., Siegers, R., Linne, M., Recker, J., Droß, P., Fräßdorf, M., Kubaty, P., Naujoks, J., Toepfer, R., Vlaeminck, S. (2018): Fachrepositorium für Forschungsdaten der Sozial- und Wirtschaftswissenschaften. Nutzungshandbuch zur Forschungsdatenbeschreibung und Datenkuration Version 0.2 (Betaphase), http://sowidatanet.de/xmlui/static/files/SDN_Nutzerhandbuch.pdf

Droß, P.J., Fräßdorf, M., Kubaty, P., Naujoks, J. (2017): Open Data in den Sozial- und Wirtschaftswissenschaften. Das Forschungsdatenrepositorium SowiDataNet. In: Jonas Kratzke/Vincent Heuveline (Hg.): E-Science-Tage 2017. Forschungsdaten managen. Heidelberg: heiBOOKS, S. 31-42. DOI: <https://doi.org/10.11588/heibooks.285.377>

Linne, M., Zenk-Möltgen, W. (2017): Strengthening institutional data management and promoting data sharing in the social and economic sciences. In: Liber Quarterly 27 (1): 58-72. DOI: <https://doi.org/10.18352/lq.10195>

Linne, M., Jensen, U. (2017): Präsentation und Live-Demo von SowiDataNet – Fachrepositorium für sozial- und wirtschaftswissenschaftliche Forschungsdaten. In: Information - Wissenschaft & Praxis 68 (2-3): 185-188. DOI: <https://doi.org/10.1515/iwp-2017-0033>

Jensen, U., Linne, M. (2017): SowiDataNet - Metadatenschema Version 1.0. GESIS Papers 2017|28. Köln: GESIS. URN: <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-54975-1>

Droß, P.J., Linne, M. (2016): Sicheres und einfaches Data Sharing mit SowiDataNet: dokumentieren - veröffentlichen - nachnutzen. In: Bibliotheksdienst. Band 50, Heft 7, Seiten 649-660, DOI: <http://dx.doi.org/10.1515/bd-2016-0079>

Droß, P. (2015): Kurzstudie: Anforderungen an die Archivierung sozial- und wirtschaftswissenschaftlicher Forschungsdaten. WZB Berlin.

Linne, M., & Klas, C. P., & Borst, T., & Toepfer, R., & Schröder, M., & Siegers, R., & Droß, P. J. (2014): Bericht zum Projektworkshop SowiDataNet: 23.06.2014, DIW Berlin.

Droß, P.J. (2014): Gesucht: Infrastruktur für Forschungsdaten. In: WZB Mitteilungen Nr. 145, S. 54-55.

Vorträge

Zenk-Möltgen, W., Linne, M., Recker, J. (2018): "Facilitating research data access – supporting data management and curation with community driven solutions." IASSIST & CARTO 2018: Once Upon a Data Point: Sustaining our Data Storytellers, 30.05.2018

Löwe, P., Linne, M., Toepfer, R. (2017): "SowiDataNet: A Research Data Infrastructure for Social Sciences and Economics". Digital Infrastructures for Research 2017, 01.12.2017

Droß, P.J., Naujoks, J. (2017): Datenkuratierung in den Sozial- und Wirtschaftswissenschaften. Qualität statt Quantität?, Open-Access-Tage, 11.-13.09.2017, Sächsische Landesbibliothek, Staats- und Universitätsbibliothek Dresden.

Droß, P.J., Fräßdorf, M., Kubaty, P., Naujoks, J. (2017): Open Data in den Sozial- und Wirtschaftswissenschaften. Das Forschungsdatenrepositorium SowiDataNet. E-Science-Tage, Universität Heidelberg, 16.-17.03.2017.

Naujoks, J. (2016): Data Sharing mit SowiDataNet. Posterpräsentation im Rahmen der internationalen Open-Access-Week in Berlin-Brandenburg vom 24. – 30. Oktober 2016, <http://doi.org/10.5281/zenodo.257721>

Toepfer, R. (2016): "Services for Storing and Sharing Research Data in Economics – Two Case Studies from Germany: ZBW Journal Data Archive and SowiDataNet" Beyond the Numbers, St. Louis (MO), USA, October 6-8, 2016

Toepfer., R. (2016): "SowiDataNet – Umgang mit Forschungsdaten in den Wirtschaftswissenschaften und potentielle Aktionsfelder für wissenschaftliche Bibliotheken" vdb Fachreferententagung für Wirtschaftswissenschaft, Frankfurt am Main, 19.-20.09.2016

Linne, M. (2016): "Improving research data sharing by addressing different scholarly target groups: individual researchers, academic institutions, or scientific journals". Embracing the 'Data Revolution': Opportunities and Challenges for Research, International Association for Social Science Information Services and Technology - IASSIST, 02.06.2016

Linne, M. (2016): "SowiDataNet – A User-Driven Repository for Data Sharing and Centralizing Research Data from the Social and Economic Sciences in Germany." Liber Annual Conference 2016: Libraries Opening Paths to Knowledge, 30.06.2016

Siegers, R. (2015, November 25./26.): SowiDataNet - Forschungsdateninfrastruktur für die Sozial- und Wirtschaftswissenschaften: Posterpräsentation beim RDA-Deutschland-Treffen, Deutsches GeoForschungsZentrum Potsdam.

Linne, M. (2015, Juli 13): SowiDataNet - Counteracting the Growing Disparity of the Data Landscape in the Social and Economic Sciences in Germany. The 6th Conference of the European Survey Research Association (ESRA) 2015.

Linne, M. (2015, Juni 2): SowiDataNet - Bringing Social and Economic Research Data Together. The 41st IASSIST Conference 2015.

Droß, P. (2015, April 21): Projekt SowiDataNet (GESIS, ZBW, DIW, WZB). SowiDataNet - Forschungsdatenrepositorium für die Sozial- und Wirtschaftswissenschaften. Posterpräsentation beim 104. Deutschen Bibliothekartag in Nürnberg 2015.

Droß, P. (2015, January 22): Projekt SowiDataNet (GESIS, ZBW, DIW, WZB). Fortbildungsveranstaltung für Fachreferentinnen und Fachreferenten der Politikwissenschaft und Soziologie an der Hertie School of Governance.

Linne, M. (2014, September 9): SowiDataNet - Aufbau eines Forschungsverbundes für Sozial- und Wirtschaftswissenschaften. Open Access Tage 2014: Session 6.

Darstellung der Maßnahmen zur Sicherung und Verfügbarmachung der im Vorhaben produzierten Forschungsdaten

Im Vorhaben wurden keine Forschungsdaten produziert.

Liste möglicher Pressemitteilungen und Medienberichte

GESIS Jahresbericht 2015, S. 22,

https://www.gesis.org/fileadmin/upload/institut/Jahresbericht_2015_Web.pdf

GESIS Jahresbericht 2013, S. 58,

https://www.gesis.org/fileadmin/upload/institut/GS_JB_2013_Web.pdf

ZBW Jahresbericht 2014, S. 31,

<https://www.zbw.eu/fileadmin/pdf/ueber-uns/jb-2014.pdf>

ZBW Jahresbericht 2015, S. 25,

<https://www.zbw.eu/fileadmin/pdf/ueber-uns/jb-2015.pdf>

ZBW Jahresbericht 2016, S. 26,

<https://www.zbw.eu/fileadmin/pdf/ueber-uns/jb-2016.pdf>

ZBW Jahresbericht 2017, S. 28,

<https://www.zbw.eu/fileadmin/pdf/ueber-uns/jb-2017.pdf>

Anhang: Überblick über die Arbeitsgruppen

AP1 Anforderungsanalyse (Datenveröffentlichung/-kuratierung/-suche)

Leitung: WZB; Partner: GESIS, DIW, ZBW

Der operative Aufbau der zu erstellenden Infrastruktur sollte auf die konkreten Bedürfnisse der Wissenschaft abgestimmt sein. Gegenstand der Anforderungsanalyse (AP1) war daher die Ermittlung von Kontexten, in denen relevante Forschungsdaten entstehen (z.B. Promotionsvorhaben, Auftragsforschung, Publikationen in Zeitschriften). Ebenso wurden Barrieren ermittelt, die einer Publikation von Forschungsdaten durch die Wissenschaftlerinnen und Wissenschaftler entgegenstehen (z.B. Urheberrechte, datenschutzrechtliche Probleme oder Restriktionen bei gekauften oder lizenzierten Datensätzen), um ihren Einfluss auf die Nutzung der zukünftigen Infrastruktur zu reduzieren.

Weiterhin wurde untersucht, welche Funktionsmerkmale die Infrastruktur aufweisen muss, um die Sichtbarmachung von Daten in anderen Kontexten zu ermöglichen (z.B. Nachweis der Daten in zentralen Katalogen, Verfügbarkeit der Daten als Linked Open Data, Vergabe von persistenten Identifikatoren für die Daten, Verknüpfung von Daten mit Publikationen). Zudem wurde geprüft, welche Community-Dienste zur Vernetzung der Datengebenden mit den Datennutzenden auf einer gemeinsamen Plattform sinnvoll angeboten werden können (z.B. Diskussionsforen, Bewertungs- oder Benachrichtigungsfunktionen).

Zentrale Outputs:

- Linne, M. u.a. (2014): Bericht zum Projektworkshop SowiDataNet. 23.06.2014, DIW Berlin.
https://sowidatanet.de/images/pdfs/Meldungen/SowiDataNet_Workshopbericht_Anforderungsanalyse_22.07.2014.pdf
- Droß, P. (2015): Kurzstudie: Anforderungen an die Archivierung sozial- und wirtschaftswissenschaftlicher Forschungsdaten. WZB Berlin.
https://sowidatanet.de/images/pdfs/Meldungen/Kurzstudie_Qualitative_Interviews.pdf

AP2 Organisationsmodell (Policies/Nutzungsbedingungen/AGBs/Institutsverträge)

Leitung: ZBW; Partner: DIW, WZB

Das Arbeitspaket Organisationsmodell (AP2) fokussierte auf das Zusammenwirken der beteiligten Einrichtungen (d.h. Repositorium, Datengebende und Datennutzende) sowie ihre jeweilige Rolle und ihren Beitrag zu SowiDataNet. Die Richtlinien zur Nutzung der Infrastruktur wurden auf Basis der Anforderungsanalyse (AP1) und in Abstimmung mit allen Beteiligten festgelegt. Dies beinhaltet unter anderem auch rechtliche Aspekte bezüglich der Gewährung von Nutzungsrechten für SowiDataNet sowie gegebenenfalls von den Nutzenden einzuräumende spezifische Rechte, die im Kontext der Langzeitarchivierung der Daten notwendig sind. Das erarbeitete Organisationsmodell wurde auf seine langfristige, über die Projektlaufzeit hinausreichende Tragfähigkeit untersucht mit dem Ziel, ein dauerhaft tragfähiges Geschäftsmodell zu erarbeiten, welches das Organisationsmodell auch finanziell nachhaltig sichert.

Teil dieses Arbeitspakets war darüber hinaus die Erarbeitung von Workflows zur Unterstützung der Forschenden bei der Speicherung und der Dokumentation von Daten in SowiDataNet in den verschiedenen Phasen des Datenlebenszyklus. Zu diesem Zweck wurde ein strukturiertes Vorgehen entwickelt, mit dessen Hilfe Fragen des Forschungsdatenmanagements durch spezialisierte Mitarbeitende in den Instituten operativ bearbeitet werden können. Exemplarische, möglichst standardisierte Workflows für die Nutzung der Infrastruktur in einem institutionellen Kontext wurden anhand konkreter Fallbeispiele entwickelt. Diese orientieren sich an den notwendigen Funktionen und Inhalten, wie sie in der Anforderungsanalyse erhoben wurden.

Zentrale Outputs:

- SowiDataNet Nutzungsbedingungen: <http://sowidatanet.de/xmlui/page/policy>
- Organisationsmodell, Darstellung in: Naujoks, J. (2016). Data Sharing mit SowiDataNet: Das Repositorium für die Dokumentation, Veröffentlichung und Nachnutzung von Forschungsdaten in den Sozial- und Wirtschaftswissenschaften <http://doi.org/10.5281/zenodo.257721>)
- SowiDataNet-Geschäftsmodell und Servicevertrag

AP3 Metadaten-Strukturen

Leitung: GESIS; Partner: WZB

Die Erarbeitung von adäquaten Metadaten-Strukturen (AP3) und die Dokumentation nach anerkannten Standards sind neben den technischen Projektkomponenten gleichermaßen wesentliche Projektbestandteile. Dabei werden sowohl aktuelle sozial- und wirtschaftswissenschaftliche Datendokumentations-Standards in Erwägung gezogen als auch relevante Standards für Bibliothekswesen und Archivierung. Ebenso werden bereits in Projektkontexten eingesetzte Schemata wie z.B. das da|ra Metadatenschema oder das DBK Metadatenschema einbezogen. In diesem Zusammenhang finden auch Schnittstellen zu Persistent-Identifizier-Systemen, dem Social und dem Semantic Web Berücksichtigung.

Zentrale Outputs:

- Jensen, U., Linne, M. (2017): SowiDataNet - Metadatenschema Version 1.0. GESIS Papers 2017|28. Köln: GESIS. <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-54975-1>

AP4 Softwareentwicklung (Repositoriumssoftware/Institutionelle Ansicht)

Leitung: GESIS; Partner: ZBW

Die Software-Entwicklung (AP4) für SowiDataNet zielte auf ein zentrales Webportal, in dem die Datengebenen die Archivierung und die Dokumentation nach anerkannten Standards durchführen können. Darüber hinaus sollten eine lokale Sicht auf die eigenen Forschungsdaten und ein differenziertes Rechte-Management ermöglicht werden. Dabei wurden die Ergebnisse der Arbeitspakete zu Anforderungsanalyse, Organisationsmodell und Metadaten berücksichtigt.

Über das SowiDataNet-Portal werden eine zentrale Suche und der Datenzugang für die Scientific Community angeboten. Zusätzlich erfolgte die Anbindung an Dienste für eine Registrierung der Daten mit dauerhaften Identifikatoren und für eine langfristige Sicherung der Daten. Die aufgebaute Infrastruktur ermöglicht eine möglichst niedrighschwellige Einspeisung lokal generierter Forschungsdaten in eine zentrale Forschungsdateninfrastruktur, wobei Fragen der Versionierung, Langzeitarchivierung, Nutzungsstatistiken und die Staffelung von Zugriffsrechten berücksichtigt wurde. Eine modulare Architektur der SowiDataNet-Infrastruktur ermöglicht die Erweiterbarkeit der Software nach der Projektlaufzeit.

Zentrale Outputs:

- SowiDataNet-Portal, <https://www.sowidatanet.de/xmlui>

AP5 Prototypentest/Beispielinhalte

Leitung: WZB; Partner: DIW

Auf Basis der Ergebnisse der Anforderungsanalyse (AP1) sowie der in AP2 entwickelten Workflows wurden durch die beteiligten Einrichtungen Beispielinhalte (AP5) für die Infrastruktur zur systematischen Generierung der entsprechenden Metadaten erarbeitet. Dabei griffen die Projektpartner auf die in ihren Instituten vorhandenen Forschungsdaten zurück. Mit diesem Vorgehen konnte die Leistungsfähigkeit der Infrastruktur für verschiedene Datentypen gezeigt, die fachliche Eignung der Metadaten geprüft und damit die Grundlage für den Weiterbetrieb nach der Projektlaufzeit geschaffen werden.

Zentrale Outputs:

- Die erprobten Workflows mit den Beispielinhalten führten zu einer Sammlung von Tickets in den Versionskontrollsystemen Mantis und GitLab, die die gefundenen Probleme und Verbesserungswünsche, sowie die Umsetzung des Bugfixing dokumentieren.

AP6 Transfer/Soziale Medien/Kommunikation

Leitung: DIW; Partner: GESIS, WZB, ZBW

Der Transfer (AP6) der entwickelten Konzepte und Infrastrukturkomponenten im Sinne einer Vermittlung an die Scientific Community wurde durch ein Schulungskonzept ermöglicht. Schulungsmaterialien wie eine ausführliche Dokumentation, Präsentationen, Broschüren und How-to-Anleitungen wurden erstellt und zentral über eine Website bereitgestellt. Außerdem wurden die Materialien über einschlägige Publikationen und die Nutzung von Social-Media-Diensten in der Community bekannt gemacht. Dazu gehörte auch das Marketing bei den Anwender-Zielgruppen und der Besuch von einschlägigen Konferenzen.

Zentrale Outputs:

- Siehe Liste der Publikationen und Vorträge
- Fecher, B. u.a. (2018): SowiDataNet - Fachrepositorium für Forschungsdaten der Sozial- und Wirtschaftswissenschaften. Nutzungshandbuch zur Forschungsdatenbeschreibung und Datenkuratation Version 0.2 (Betaphase), http://sowidatanet.de/xmlui/static/files/SDN_Nutzerhandbuch.pdf