




Herausforderungen der Datenintegration (Erhebung, Verknüpfung und langfristige Sicherung) in der Forschungsprojekten an Hochschulen

Prof. Dr. Cordula Artelt

Herrenhäuser Konferenz Datenräume in Deutschland und Europa gestalten –
Impulse der Wissenschaft | 24.04.2023



Herausforderungen der Datenintegration (Erhebung, Verknüpfung und langfristige Sicherung) in der sozialwissenschaftlichen Forschung

WARUM WIR DATENINTEGRATION UND NEUE DATENRÄUME BRAUCHEN

Survey Daten: Large-Scale Panel oder Trend Studien, z.B.

SOEP | NEPS | FReDA | FGZ | PASS | ALLBUS | ESS | PISA | TIMMS | PIAAC ..

Stärken: z.B. bekannte Stichprobenbasis | psychometrische Gütekriterien

Georeferenzierte Daten

...

Big data

...

Registerdaten

...

Synthetische Daten

...

Neue Arten von Datenerhebungen, z.B.

Virtual reality | NLP | Augmented reality | Ecological momentary assessment

Stärken: z.B. Agilität | Dynamik | Dichte

Stärken der Integration und Synergien

- Validitätsgewinne & neue Konstrukte
- Reichere Daten & Effizienz

BEISPIEL NEPS

FILME STATT MOMENTAUFNAHMEN



> ZIEL

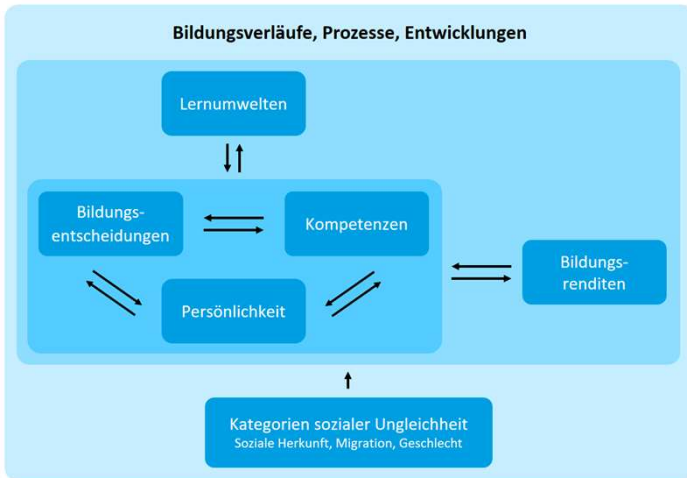
Aufgabe des Nationalen Bildungspanels ist es, **qualitativ hochwertige Daten** zu Bildungsprozessen und der Entwicklung von Kompetenzen über den Lebenslauf bereitzustellen.

> KONZEPTION

Das NEPS ist als **Längsschnittstudie** konzipiert, bei der die Befragten im Bildungsverlauf begleitet werden (im Gegensatz zu Querschnittstudien wie z.B. PISA). Es umfasst **alle Bildungsphasen im Lebenslauf** – von der frühen Kindheit bis ins hohe Erwachsenenalter.

> #DATASHARING

Alle erhobenen Daten werden der **Wissenschaftsgemeinschaft zeitnah in nutzerfreundlicher Form** mit modernsten technischen Systemen **zugänglich gemacht**.



Über **70.000 Personen** vom Kleinkind bis zum Erwachsenen wurden und werden im NEPS befragt und getestet.

Dazu kommen mehr als **50.000 Personen** aus dem **Umfeld** der Befragten, z.B. Eltern oder pädagogisches Fachpersonal.



Mehr als **140+ Haupterhebungen** wurden seit Beginn des NEPS durchgeführt, über **500+ Befragungs- und Testinstrumente** wurden dabei eingesetzt.

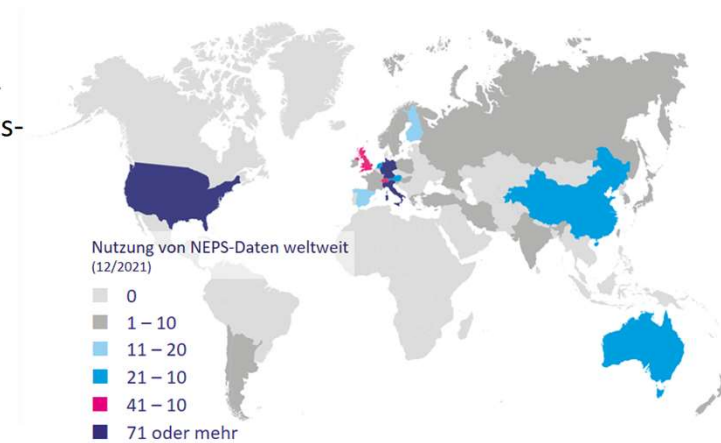


90+ Scientific-Use-Files* wurden seit 2011 veröffentlicht – sie stehen der Wissenschaftsgemeinschaft kostenfrei zur Verfügung.

* inkl. Zusatzstudien Thüringen und Baden-Württemberg



Mehr als **3.000+ Forschende** arbeiten bereits mit NEPS-Daten – in **32 Ländern** weltweit.



NEPS: STÄRKEN UND LIMITATIONEN DES NEPS

▪ **Stärke des NEPS: komplexe Daten für die Bildungsforschung**

- Reichhaltige Informationen über multiple Bildungskontexte, Entwicklungsdimensionen, subjektive Wahrnehmungen, Einstellungen, Entscheidungen und Outcomes im Längsschnitt
- Erklärung von Bildungsentscheidungen, -verläufen, Kompetenzentwicklung und Ergebnissen

▪ **Limitation des NEPS, z.B. Kontextdaten**

▪ **Erfassung von Kontextdaten**

- Fast alle Kontextinformationen im NEPS werden bislang erfragt.

▪ **NEPS-Kontextdaten sind lückenhaft**

- Welche Merkmale, Ressourcen, Programme haben die NEPS-Schulen?
- Wie sieht ihre räumliche Lage, Infrastruktur und Nachbarschaft aus?
- Wie sind Schulen und Klassen zusammengesetzt?

▪ **Grenzen dieser Kontexterfassung**

- Individuen wechseln Schulen, NEPS verliert den institutionellen Kontext

Amtliche Daten

Georeferenzierte Daten

Bildungsverlaufsregister

HERAUSFORDERUNGEN UND HANDLUNGSBEDARF

- Surveys bieten reichhaltige Informationen auf Basis robuster Designs; sie stehen aber vor Herausforderungen:
 - Geringe Effizienz, auch und insb. in der Erfassung von Kontexten
 - Lange Befragungs- und Testdauer | Hohe Belastung der Teilnehmenden
 - Zurückgehende Teilnahmebereitschaft in Haushalten und Institutionen
 - Sehr hohe Kosten | Steigende Preise von Umfrageinstituten
 - Nicht-Nutzung der Potenziale integrierter und verknüpfter Daten

“Not getting behind” – New Data landscapes als Desiderat

Mission: Understand and push forward issues specific to new data types and their linkage from different perspectives and develop strategies for incorporating these data into different types of robust designs of social sciences

Warum Potenziale (noch) nicht genutzt werden (können)

- Verfügbarkeit von Daten
- Rechtliche und politische Rahmenbedingungen
- Tools und *know-how* für die Integration und Nutzung verschiedener Erhebungs- und Datenarten



VISION: NEPS ALS BESTANDTEIL INTEGRIERTER BILDUNGSDATEN

Schulstatistik

- Größe
- Räumliche Umgebung
- Zweige, Schwerpunkte
- Programmteilnahme
- Ressourcen

Lernstand / Kompetenzen

- VERA
- IQB-Bildungstrends
- NEPS

Längsschnittsurveys (NEPS)

- Wohlbefinden
- Wahrnehmungen
- Sozio-emotionale
Kompetenzen
- Persönlichkeit
- Interessen
- Aspirationen
- Entscheidungen
- Integration, Politik,
Gesundheit
- Weiterbildung
- Soziales Umfeld

Bildungsverlaufsregister

- Hintergrundmerkmale
- Räumliche Umgebung
- Schul-, Ausbildungs-, Studienverlauf
- Zweige, Schwerpunkte, Fächer
- Noten und Abschlüsse
- Zusammensetzung von Schulen

Schlüssel

- Person
- Schulen

Andere Register

- Soziale Herkunft
- Erwerbsbiographien
- Einkommen

- Vorbild: europäische Nachbarländer wie Niederlande, Dänemark, Österreich, Italien

INFRASTRUCTURE PRIORITY PROGRAMME „NEW DATA SPACES FOR THE SOCIAL SCIENCES“

Deutsche Forschungsgemeinschaft (German Research Foundation)
Information for Researchers

Call for Proposals

No. 20

14 March 2023

Infrastructure Priority Programme “New Data Spaces for the Social Sciences” (SPP 2431)

In December 2022, the Senate of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) established the Infrastructure Priority Programme (InfPP) “New Data Spaces for the Social Sciences” (SPP 2431) as a long-term funding scheme. The programme will initially run for six years (the first funding period). If it is successful, the programme committee will consider applying for an additional six years. The present call invites researchers to submit proposals for the first three years of the first funding period.

Challenge



Contact Home anmelden

Infrastructure Priority Programme Call Workshop Downloads

Call for Proposals

Applications for project funding can be submitted until mid-September 2023!

[To the Call](#)



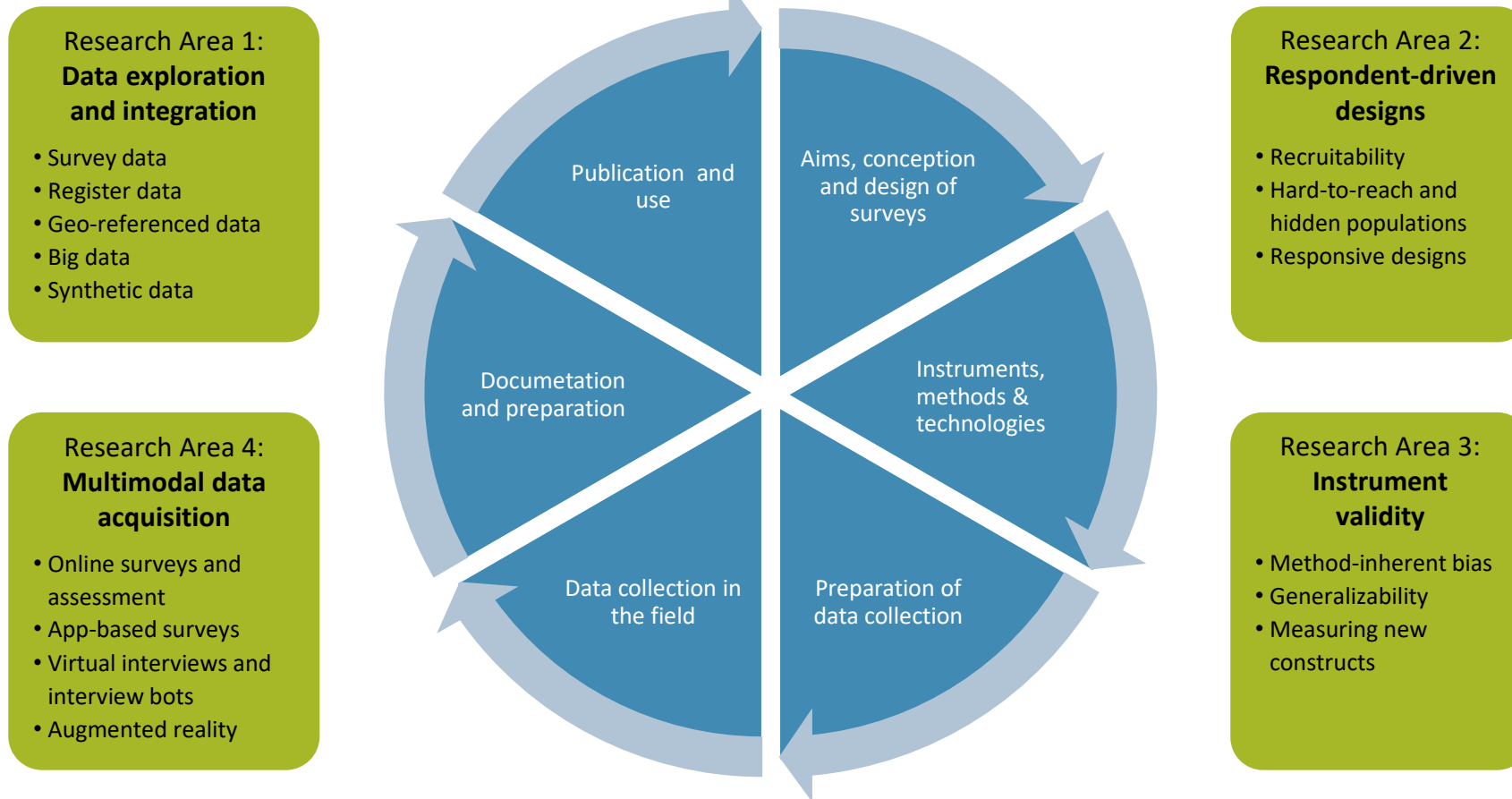
DFG Infrastructure Priority Program (SPP 2431)

NEW DATA SPACES FOR THE SOCIAL SCIENCES

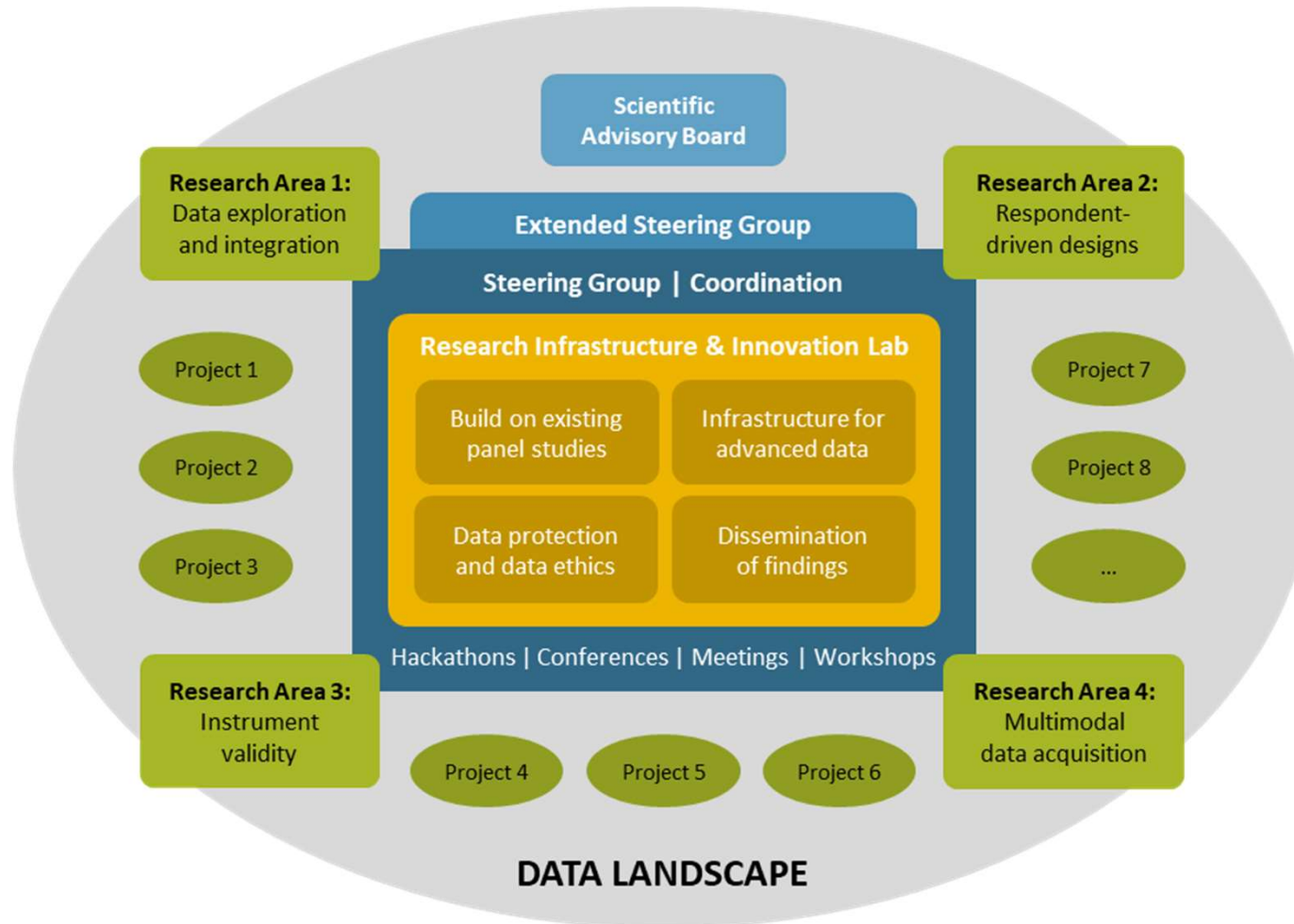
The priority program “New Data Spaces for the Social Sciences” aims to open up new data spaces for the social sciences. It is designed for the long term and, in view of current social

www.new-data-spaces.de

SURVEY LIFE-CYCLE UND EMERGING FIELDS IN DER FORSCHUNG



NEW DATA SPACES FOR THE SOCIAL SCIENCES



UNLOCK OPPORTUNITIES OF NEW DATA SPACES



- Systematische und weitsichtige Forschung in den Sozialwissenschaften muss sich bemühen um
 - **Neue Datenformen** | z.B. Daten aus administrativen Prozessen, Digitale Verhaltensspuren
 - **Neue Formen der Datenerhebung** | z.B. virtuelle Interviews und augmented reality
 - **Neue Qualität der Messung von methoden-inherentem Bias** | z.B. machine learning, KI vs. self-report



**LEIBNIZ-INSTITUT FÜR
BILDUNGSVERLÄUFE**

Wilhelmsplatz 3
96047 Bamberg
www.lifbi.de

NEPS-Entwicklung

Telefon: +49 951 863-3433
neps-development@lifbi.de



twitter.com/LifBi_Bamberg

