

DIGITALISATE ZWISCHEN ERKLÄREN UND VERSTEHEN

Chancen und Herausforderungen durch Big Data für die Kultur- und Sozialwissenschaften – Eine wissenschaftstheoretische Desillusionierung

(Dr. Gransche & Prof. Gethmann)

Berlin, 18.10.2018

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Das Projekt „ABIDA – Assessing Big Data, Big Data Begleitforschung“ wird vom BMBF gefördert (Förderkennzeichen 01|S15016A-F)

GEGENSTAND

Gegenstand

Frage: Relevanz von Big Data für die Geisteswissenschaften und die verstehenden Sozialwissenschaften?

Möglichkeiten: Die Forschung in der jeweiligen Disziplin ist

- a) **bereichert** durch neue Methoden basierend auf Big Data,
- b) **herausgefordert** mit möglichen Nachteilen durch das Aufkommen von Big Data (z. B. neue Kompetenzen oder erhöhter Rechtfertigungsdruck) umzugehen, oder
- c) *de facto* von Big Data (u. U. im Gegensatz zur Debattenlage) gar **nicht betroffen**.

Antworten je für Einzelbereiche von Soziologie,
Politikwissenschaft, Ökonomie, historische und philologische
Disziplinen etc....

...kann ein Einzelgutachten nicht geben.

Das Gutachten widmet sich vielmehr:

- den **epistemologischen und methodologischen Kernfragen**,
- die aus **wissenschaftstheoretischer und technikphilosophischer Sicht**
- für die Frage nach den **Herausforderungen** durch Big Data-Technologien für die Geistes- und verstehenden Sozialwissenschaften hohe Relevanz haben.

Inhalt

1. Überblick über den Untersuchungsgegenstand
2. Schutz der Individualität vor Stratifizierung
3. Korrelationsanalyse und Kausalerklärung
4. Big Data und Vorhersagen

Überblick, Abgrenzungen, vorbereitende Klärungen

KAPITEL 1

Kapitel 1

1. **Überblick über den Untersuchungsgegenstand**
 - 1.1 Wissenschaftsklassifikationen
 - 1.2 Natur – Geist
 - 1.3 Objektiver Geist - Kultur
 - 1.4 Verstehen und Sinn
 - 1.4.1. Verstehen I: Kritik der „expressionistischen“ Hermeneutik
 - 1.4.2. Verstehen II: „konventionalistische“ Hermeneutik
 - 1.5 Erklärende und verstehende Sozialwissenschaften
 - 1.6 Big Data und die verstehenden Wissenschaften

Epistemologische Kernfrage

- These im Zuge von Big Data:
 - Herausforderung für Geistes-/Kultur- und Sozialwissenschaften
 - Hoffnung, direkt von den Daten relativ interpretationsfrei Zusammenhänge „aus- und ablesen“ zu können, ohne auf als fehlbar angesehe Auslegungsinstanzen rekurrieren zu müssen.
- Leitdifferenz: Erklären – Verstehen
 - Erklärende Wissenschaften (Natur-) – *sciences*
 - Verstehende Wissenschaften (Geistes-, Kultur-, Sozial-) – *humanities*

Methodologische Kernfrage

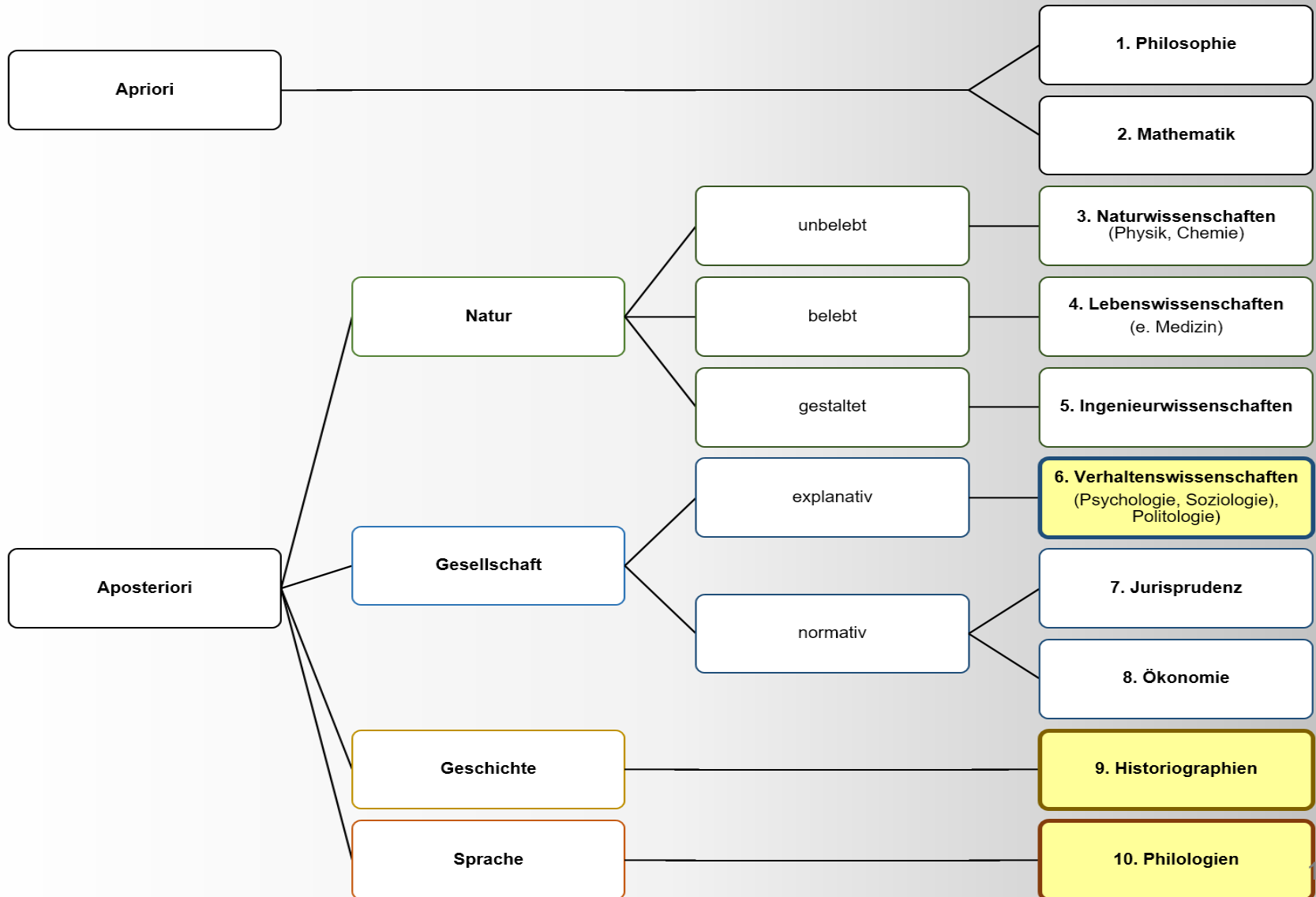
- Nicht: „Welche Wissenschaften sind Geisteswissenschaften?“ (und die Antwort wäre eine Liste solcher Wissenschaften...), sondern „Welche Erkenntnisverfahren (Methoden) sind ‚geisteswissenschaftlich‘, d.h. verwenden einen *so und so gearteten* Verstehensbegriff“ (und die Antwort ist eine Liste von Erkenntnisverfahren).
 - Formen des Verstehens in (fast) allen Disziplinen: *einen Text verstehen* auch Bauingenieure, *eine Formel verstehen* auch Linguisten, *ein Röntgenbild verstehen* Mediziner und Archäologen, *die Veränderung eines Organismus verstehen* Biologen...
 - Unterscheidung von verstehenden und erklärenden Wissenschaften durch Rekonstruktion der Verstehenselemente
 - rekonstruierende Lehre vom Verstehen heißt traditionell „Hermeneutik“

Verstehende Wissenschaften?

- Geisteswissenschaften als Wissenschaften
 1. von (Geist als) *Naturgegenständen* (z. B. Neurowissenschaften)?
 2. vom subjektiven *Geist* (z. B. Psychologie)
„Form der Beziehung auf sich selbst“ (Hegel)
 3. vom **objektiven *Geist*** mit Verbindlichkeits- und Anerkennungsstruktur
„Form der Realität als einer von ihm hervorzubringenden und hervorgebrachten Welt, in welcher Freiheit als vorhandene Notwendigkeit ist“ (Hegel) → also von Menschen gemacht, dem Individuum vorgegeben
z. B.: Sprache, Geld, Sitten, Staat, Recht, Kunst, Wissenschaft etc. → Kultur

“It is difficult to translate the term *Geist* and impossible to indicate succinctly in English relationships between words that are obvious in German.” *The Encyclopedia of Philosophy*

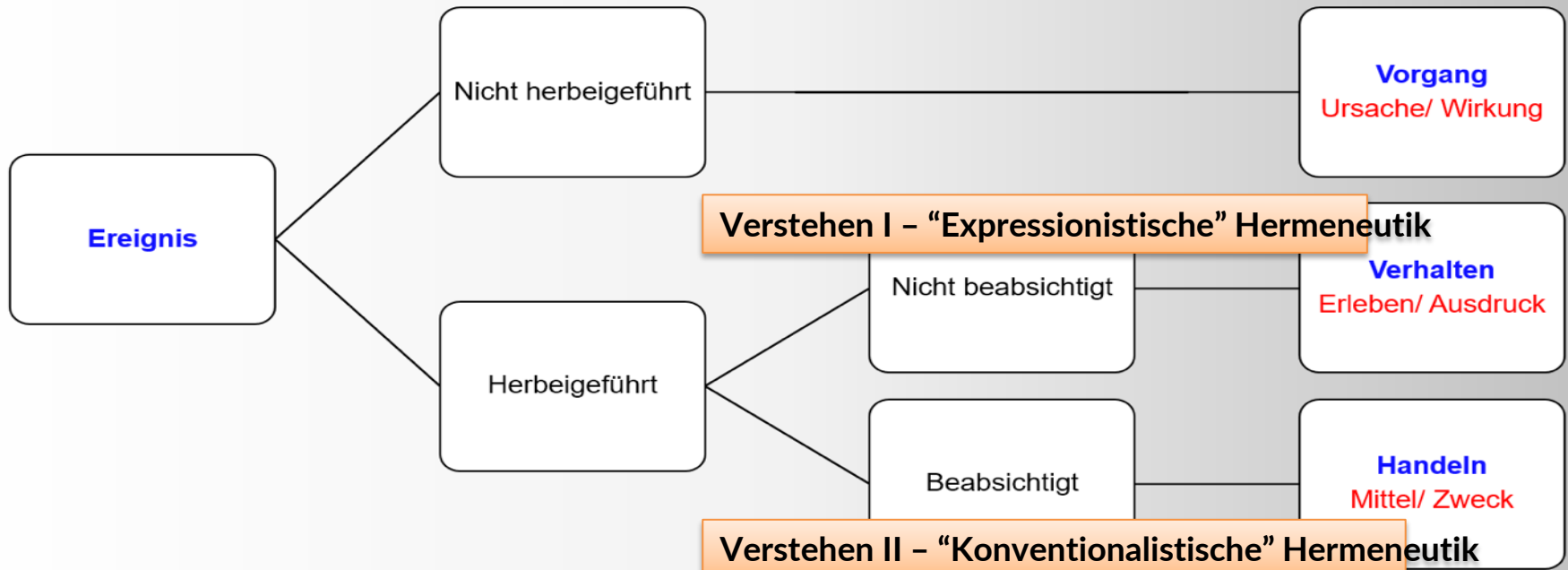
Kultur- /Geistes- und Sozialwissenschaften?



Kleine Ontologie der Wissenschaften

Kleine Ontologie der Wissenschaften

Gegenstände / Deutungskategorien



Big Data & Kapitel 1

- Wären die Digitalisate des Datenschattens oder allgemein die algorithmisch durchforstete Datenstruktur überhaupt als Spuren von Verhalten, also Zeichen eines Ausdrucks zu verstehen? Ist ein GPS-Profil eine Verhaltens-, gar eine Handlungs- oder nicht vielmehr eine Vorgangsobservable?
- Mit der zunehmenden Sensorverbreitung, der Digitalisierung und mit Big Data-Anwendungen werden zunächst nur mehr/andere *Vorgänge* erfassbar, speicherbar und vergleichbar, *Handlungen* jedoch keinesfalls.

Big Data & Kapitel 1

Beispiel:

- Der Versuch bspw. mittels emotionssensitiver Technologien das Erleben unter Umgehung verstehender Instanzen „auszulesen“ (Der Autofahrer ist vom Reh überrascht worden.), greift erkennbar zu kurz.
- Das Big Data gestützte Zurückführen eines Ausdrucks (in Gestalt von algorithmisch erkennbarer Muster wie bspw. Mimik) auf ein Erleben (z. B. Überraschung) verirrt sich in den Reduktionsverhältnissen.
- Sensorisch und algorithmisch erfassbar sind einzig Vorgänge (z. B. Bewegungen von markierten Gesichtspunkten). Die Verknüpfung dieser Vorgänge mit Verhaltensmustern, Erlebenstypen oder gar die Zuordnung von Handlungsintentionen ist und bleibt der Big Data-Anwendung extern bzw. vorgegeben.

Big Data & verstehende Wissenschaften

- **Erklärung:** Rückführung eines Zustandes oder Ereignisses auf seine Gründe oder Ursachen. Erklärungszusammenhänge können deterministisch oder probabilistisch voneinander abhängen (lediglich bei letzteren spielen große vernetzte Datenmengen (Big Data) eine relevante Rolle).
- **Verstehen:** richtet sich auf Verallgemeinerbarkeitsansprüche, für deren Verständnis große vernetzte Datenmengen (Big Data) nicht konstitutiv sind. Aber:
 - algorithmische Strukturvorschläge/ Big Data-Erklärungsvorschläge
 - als Hilfe bei der Hypothesenbildung
 - als Kreativitätstechnik im Prozess der Erstellung von Verallgemeinerbarkeitsansprüchen
 - primär abduktives Schließen (unsicher, kreativ)

Individuum und Datenprofil

KAPITEL 2

Kapitel 2

- 2. **Schutz der Individualität vor Stratifizierung**
- 2.1 Wissenschaftstheoretische Grundlagen von Big Data
- 2.2 Big Data-Industrie als Treiber der Stratifizierung
- 2.3 Individualität und Exemplarität
- 2.4 Vollindividuum und Datenschatten

- **Big Data**

- Mustererkennung, Korrelationen
- Treiber der Stratifizierung (bspw. Scoring)
- Individuen als Exemplare
- IT-technische Verarbeitung – und sei es im Big Data Ausmaß –, bleibt aus prinzipiellen Gründen auf die Seite **Exemplarität** beschränkt.

- **Verstehende Wissenschaften**

- Geistes- und verstehenden Sozialwissenschaften (u.a.) Individuum nicht nur Exemplar einer Spezies (Element einer Menge)
- Gegenstände des objektiven Geistes (und damit der Geisteswissenschaften) sind **Individuen**, nicht Exemplare für etwas.

Vollindividuum vs. Datenschatten

Erkenntnisinteresse
diese sollen verstanden werden
über diese soll etwas erfahren werden

Erkenntnisquelle
anhand dieser soll verstanden werden
von diesen wird erfahren

realem Subjekt
(Voll-)Individuum

virtuellem Subjekt

Datenschatten

Nutzer

(Nutzer-)Profil

Individuum

stratifizierter Kohorte

Individualität

Exemplarität

Verhalten, Handeln

Vorgänge

Ursache, Erleben, Zweck

Wirkung (Ausdruck, Mittel)

Bsp.: Jogging-Aktivität

Bsp.: GPS-Daten

Unüberbrückbare Kluft

- Die durch Stratifizierung vorgenommene Generalisierung verhält sich antagonistisch zur Methode des Verstehens der Gegenstände des objektiven Geistes.
- Die Kluft kann aber je nach Anwendung
 - **klein genug** sein (z.B. für Entscheidungen im Shop Floor-Management oder Predictive Pricing) oder
 - **sogar zu klein** (Verwechslungsgefahr durch die Individuen).

Kausalität und Korrelation

KAPITEL 3

Kapitel 3

- 3. **Korrelationsanalyse und Kausalerklärung**
- 3.1 Determinismus und Probabilismus
- 3.2 Pseudo-Erklärungen
- 3.3 Wissenschaftsphilosophische Rekonstruktionsansätze des Kausalbegriffs
- 3.4 Korrelation und Kausalität in den Geistes- und verstehenden Sozialwissenschaften

Korrelation und Kausalität in den Geistes- und verstehenden Sozialwissenschaften

- Sowohl probabilistische als auch deterministische **Kausalanalysen** sind auf den Bereich beschränkt, der durch den Begriff der **Erklärung** charakterisiert ist (durch Gesetz oder Wahrscheinlichkeit).
- **Geisteswissenschaften** (und einige Theorieansätze der Sozialwissenschaften) sind jedoch auf die Erkenntnisdomäne des **Verstehens** bezogen.
- Da **Korrelationsanalysen** immer, unabhängig davon ob mit großen vernetzten Datenmengen gearbeitet wird oder nur mit kleinen Datensammlungen, auf Erklärungen abheben, sind sie für eine recht verstandene geisteswissenschaftliche Methode nicht „disruptiv herausfordernd“ einschlägig (damit jedoch keineswegs nutzlos).

Big Data und Vorhersagen

KAPITEL 4

Kapitel 4

- 4. **Big Data und Vorhersagen**
- 4.1 Vorhersagen versus datengestützte Konditionalaussagen
- 4.2 Vorhersageabhängigkeit
- 4.3 Die Handlungsorientierung
- 4.4 Möglichkeit der Vorhersage
- 4.5 Big Data und Vorhersagen?
- 4.6 Big Data als Krakenorakel
- 4.7 Konsequenzen für die Sozial- und Geistes-
/Kulturwissenschaften

Vorhersagen \neq Konditionalaussage

- „Korrekte“ Vorhersagen sind unmöglich (kein Wahrheitskriterium für Zukünfte)
- Aussagen über Kausalitäten und Wahrscheinlichkeiten sind nützlich und riskant zugleich
 - „Wenn A, dann B.“ = Konditionalaussage („immer...“ \rightarrow kausal)
 - „Übermorgen wird B.“ = Vorhersage
 - Problem: Wann ist A gegeben? (s. angewandte Ethik)
- Feststellung des Gegebenen kann durch Big Data profitieren

„Wie ist eine Geschichte a priori möglich? Antwort: wenn der Wahrsager die Begebenheiten selber macht und veranstaltet, die er zum Voraus verkündet.“ Kant, *Der Streit der Fakultäten*, 1798, II.2.

Big Data & Vorhersagen

- Vorhersage(un)abhängige Systeme (va-S, vua-S)
- vua-S: z. B. Planetenbahnen | va-S: z. B. Börse
- Vorhersageabhängigkeit, Rezeptivität für den Sprechakt, va-S sind sinnverarbeitende Systeme (z. B. Sprache verstehend) → Verstehende Systeme
- Big Data Vorhersagen wirken in va-S zukunftsgestaltend (wenn auch schwerlich voll intentional)
- Die Wirkung (Mächtigkeit der Handlungsorientierung) hängt von der Aufnahme im va-S ab also von der Verbreitung, Inszenierung, Relevanz, Vertrauen/Akzeptanz etc.
- Big Data als Hype/ Verheißung/ Herausforderung/ Bedrohung mit großer gesellschaftlicher Relevanz (auch zur Rechtfertigung von ABIDA selbst)

Konsequenzen für die verstehenden Wissenschaften allg.

- Für die (sozialwissenschaftlichen) Erhebungsmethoden bedeutet Big Data eine **Chance**: „Reduktion“/ Transformation der Beobachtungsstörung (Big Data als „omnipräsenter nicht mehr wahrgenommener digitaler Spionsspiegel“)
- Es geht nicht darum, ob datengestützte Erhebungen eine interferenz*freie* Beobachterinstanz sein könnten (können sie nicht), ...sondern darum, ob manche Phänomene sich im Vergleich zu bislang zur Verfügung stehenden Methoden interferenz*ärmer* beobachten lassen.
- Probleme: a) der Untersuchungsgegenstand der Sozialwissenschaften zählt zu den va-Systemen, b) „Unschärferelation“

Konsequenzen für die verstehenden Wissenschaften – Vorhersage

- Tendenz der **Personalisierung** durch lernende IT-Systeme, die wiederum durch Big Data ermöglicht werden, stellt eine **Typenfokussierung** über eine Individuenfokussierung (Kapitel 2).
- **Individuen** laufen Gefahr, in anonymer Vergemeinschaftung, als vorausberechenbares **Stratus-Exemplar** auf **erwartbare Entwicklungen reduziert** zu werden; dies kann – zumal in einer umfassend technisierten Gesellschaft – zu weitreichenden Problemen für die Würde und Autonomie des Individuums führen.

Konsequenzen für die verstehenden Wissenschaften – Vorhersage

- Gefahr, dass der Kategorienfehler (einer Verwechslung von Vorgang, Verhalten und Handeln) sich durch die technische Zurichtung und Vorhersageausrichtung des Alltags ex post als doch zutreffend herausstellen könnte (unser Handeln und Verhalten auf die verknüpften Vorgänge reduziert wird).
- Durch Big Data-basierte Systeme die Berechenbarkeit virtueller Subjekte über die Entscheidungsautonomie realer Subjekte zu stellen, wäre als **vorhersagebasierter Paternalismus** zurückzuweisen (praktikable Anwendungsbereiche bewusster Zulassung zugestanden).

ZUSAMMENFASSUNG

Zusammenfassung

- Big Data richtet sich
 - primär auf Wahrscheinlichkeit also auf Erklären, weniger auf Verstehen
 - damit eher auf (erklärende) Naturwissenschaften bzw. datenverarbeitende Wissenschaften (v.a. 3, 4 und 5)
 - auf Exemplare von Strata, Datenprofile, nicht auf Individuen

Zusammenfassung

- Big Data-Technologien könnten das notorische Problem der geringen Fallzahlen seitens der empirischen Sozialforschung relativieren helfen
- Big Data-Funde könnten zum Gegenstand in den verstehenden Wissenschaften gemacht werden
- Big Data könnte hilfreich (nicht ersetzend) für die Hypothesenbildung, Ideengenerierung, als Kreativitätstechnik...

Zusammenfassung

- Dass eine unüberbrückbare Kluft (zwischen Individuum und Datenschatten) besteht, was bspw. einer Big Data basierten Soziologie vorgehalten werden könnte, bedeutet nicht, dass diese Kluft nicht gf. immerhin „kleiner“ ist, als die der bisherigen Methoden.
- Big Data Anwendungen könnten dort helfen, wo Erklärungsanteile in den verstehenden Wissenschaften relevant werden – vice versa stoßen sie selbst in den erklärenden Wissenschaften dort an Grenzen, wo diese Verstehensanteile aufweisen.

Big Data: Relevanz für verstehende Wissenschaften?

bereichert	herausgefordert	nicht betroffen
<ul style="list-style-type: none">• Hilfsmittel• Kreativitätstechnik• Aufmerksamkeitsinstrument• u.U. mit interferenzärmeren Beobachtungsmittel	v.a. im Verstehen der Konsequenzen von Big Data (Stratifizierung etc.) (ABIDA), weniger als “obsolete” Methoden/ Wissenschaften	im Verstehen (vs. Erklären), im Individuellen (vs. Exemplarischen)
<ul style="list-style-type: none">• nicht mit “objektiven” Daten,• nicht mit interpretationsfreiem/ theorielosem Auslesen der Welt,• nicht mit interferenzfreien Beobachtungen	Big Data kann kein Verstehen ersetzen also auch nicht die verstehenden Wissenschaften obsolet machen: das kann nur die Gesellschaft, die Politik und die Wissenschaften selbst	

**WEITERE
FORSCHUNG
NÖTIG!**

Vielen Dank für Ihre Aufmerksamkeit!

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung