

Interdisziplinäre Konferenz des Rates für Informationsinfrastrukturen (RfII)  
mit Unterstützung der VolkswagenStiftung

## HERAUSFORDERUNG DATENQUALITÄT – FORSCHUNG IM DIGITALEN WANDEL

#dataQ2020

27. bis 28. Februar 2020

Xplanatorium Schloss Herrenhausen, Hannover

Die Sicherstellung von Datenqualität im Zuge des digitalen Wandels berührt den Kern wissenschaftlichen Arbeitens. Dass Fragen der Datenqualität heute alle Akteure der Forschung, der Infrastrukturen und der Wissenschaftspolitik vor neue Herausforderungen stellen, betont der Rat für Informationsinfrastrukturen (RfII) in seinem Positionspapier „Herausforderung Datenqualität – Empfehlungen zur Zukunftsfähigkeit von Forschung im digitalen Wandel“.

Ziel der Konferenz ist es, diesen Herausforderungen diskursiv zu begegnen. Hierzu werden neben Akteuren aus Wissenschaftspolitik und Forschungsförderung Vertreterinnen und Vertreter verschiedenster wissenschaftlicher Disziplinen Beiträge beisteuern und mit den Teilnehmerinnen und Teilnehmern diskutieren.

### PROGRAMM

DONNERSTAG, 27. FEBRUAR 2020

12:00 Uhr    Snack und Registrierung

13:00 Uhr    Grußworte

- DR. GEORG SCHÜTTE, Generalsekretär der VolkswagenStiftung, Hannover
- PROF. DR. STEFAN LIEBIG, Stellvertretender Vorsitzender des Rates für Informationsinfrastrukturen (RfII)
- DR. MICHAEL MEISTER, Parlamentarischer Staatssekretär im Bundesministerium für Bildung und Forschung (BMBF)

**13:40 Uhr    Auftakt: Datenqualität als wissenschaftspolitische Herausforderung**

Gut dokumentierte und interoperable Forschungsdaten sind eine wichtige Voraussetzung für den Erkenntnisfortschritt der Wissenschaft im digitalen Wandel. Sie leisten darüber hinaus einen wesentlichen Beitrag zum Wissenstransfer in Innovationsprozessen und zur Erhaltung des Vertrauens der Öffentlichkeit in die wissenschaftliche Forschung. In der Auftaktveranstaltung werden hierauf bezogene Impulse aus der Perspektive der Wissenschaftsorganisationen und der Wissenschaftspolitik gegeben.

Moderation:

- SABINE BRÜNGER-WEILANDT, Direktorin und Geschäftsführerin des Leibniz-Instituts für Informationsinfrastruktur (FIZ), Mitglied des RfII

Keynotes:

- PROF. DR. MARTINA BROCKMEIER, Vorsitzende des Wissenschaftsrates (WR) 2017 – 2020:

*Welchen Kulturwandel braucht das Wissenschaftssystem, wenn Forschung datenintensiver wird?*

- PROF. DR. PETRA GEHRING, Vorsitzende des Rates für Informationsinfrastrukturen (RfII):

*Datenqualität gestalten: Notwendige Rahmenbedingungen für gute wissenschaftliche Praxis im digitalen Wandel*

- PROF. DR. RALF B. WEHRSPORN, Vorstand für den Bereich Technologiemarketing und Geschäftsmodelle, Fraunhofer Gesellschaft (FhG):

*Forschungsdaten für Innovationsprozesse: Optionen für Anreizsysteme – Erwartungen an die Qualitätssicherung*

**15:00 Uhr    Pause****15:30 Uhr****Session I: Datenqualität in Simulationsstudien, Modellierungen und maschinellem Lernen**

Simulationen und Modellierungen sind voraussetzungsvolle Forschungsprozesse und stellen sowohl auf der Ebene der Selektion und Einspeisung „externer“ Daten als auch hinsichtlich der auf dieser Basis „intern“ erzeugten Daten hohe Qualitätsanforderungen. Auch Erfolge im maschinellen Lernen sind stark von der Qualität der verwendeten Trainingsdaten abhängig. Gemeinsam ist diesen Forschungsformen das Interesse, die Validität der Daten und die Transparenz der Datenerzeugung zu optimieren sowie „Daten Bias“ und hieraus resultierende Pfadabhängigkeiten nach Möglichkeit zu eliminieren. Wie dies gelingen kann und wo weitere Herausforderungen liegen, wird beispielhaft aufgezeigt.

Moderation:

- PROF. DR. WOLFGANG MARQUARDT, Vorstandsvorsitzender des Forschungszentrums Jülich, Mitglied des RfII

Impulsvorträge:

- PROF. DR. BODO ROSENHAHN, Institut für Informationsverarbeitung, Leibniz Universität Hannover:

*Big Data = Good Data? Modellbildung im Angesicht von unbalancierten, verzerrten und domänenspezifischen Daten*

- PROF. DR. GABRIELE GRAMELSBERGER, Wissenschaftstheorie und Technikphilosophie, RWTH Aachen:

*Selbstlernende Klimamodelle. Modellierung, maschinelles Lernen und Evaluation von Daten*

- PROF. DR. PETER PELZ, Institut für Fluidsystemtechnik, Technische Universität Darmstadt:

*Datenqualität als Herausforderung für die Modellierung von Infrastrukturen für hyperurbane Regionen*

17:00 Uhr      Pause

17:30 Uhr      Evening Lecture (mit anschließender Diskussion)

- PROF. DR. ROLF-DIETER HEUER, Senator der Helmholtz-Gemeinschaft für den Forschungsbereich "Materie" und ehemaliger Generaldirektor des CERN:

*Datenqualität und Datenmenge als Spannungsfeld in großen Forschungsinfrastrukturen: Was wir aus den Erfahrungen des CERN lernen können*

19:00 Uhr      Apéritif und gemeinsames Abendessen der Konferenzteilnehmer (Buffet)

FREITAG, 28. FEBRUAR 2020

9:00 Uhr

**Session II: Qualitätsvolle Daten aus dem Internet (?) – Chancen und Herausforderungen für die Forschung mit Sozial- und Wirtschaftsdaten**

Im Zuge der Digitalisierung stehen konventionelle methoden- und theoriebasierte Forschungsansätze der Survey-Forschung in einem starken Wettbewerb mit der hands on-Datenanalyse in sozialen Medien und Plattformen im Internet. Für die Weiterentwicklung der Sozial- und Wirtschaftswissenschaften ist es von vitalem Interesse herauszufinden, ob und wie die Forschung mit Daten aus dem Internet mit akzeptierten Methoden und Datenqualitätsstandards verknüpft werden und "konventionelle" Erhebungen ergänzen kann.

Moderation:

- PROF. DR. STEFAN LIEBIG, Direktor des Sozio-oekonomischen Panels (SOEP) und Vorstandsmitglied des Deutschen Instituts für Wirtschaftsforschung (DIW Berlin), stellvertretender Vorsitzender des RfII

Impulsvorträge:

- PROF. JULIA INGRID LANE, PhD, Wagner School for Public Policy, New York University:

*Building New Data Infrastructures: A Practical Approach*

- PROF. DR. FRAUKE KREUTER, Statistik und Methoden, Universität Mannheim und Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit:

*Data Collection in the Aera of Digitalisation: Understanding Data Generating Mechanisms*

- PROF. JANA DIESNER, PhD, School of Information Science, University of Illinois at Urbana-Champaign:

*Reliable Signals? Discovering the Impact of the Quality of Social Interaction Data on Social Science Theory, Knowledge, and Practical Applications*

10:30 Uhr    Pause

11:00 Uhr

**Session III: Von der informellen „guten Praxis“ zur Data Governance**

In Wirtschaftsunternehmen haben sich Ansätze einer "Data Governance" entwickelt, um Verantwortlichkeiten und formalisierte Prozesse für den Umgang mit Daten zu definieren sowie die Verlässlichkeit datenbasierter Entscheidungen zu gewährleisten. In der Wissenschaft fungieren Bereiche als Schrittmacher, für die gelingender Transfer von Forschungsdaten und die Verknüpfung mit Daten Dritter von großer Bedeutung sind. Hierzu zählen u.a. die Lebenswissenschaften, die klinische Forschung und die organisationsbezogene Innovationsforschung.

Moderation:

- MARIT HANSEN, Landesbeauftragte für Datenschutz Schleswig-Holstein, Mitglied des RfII

Impulsvorträge:

- PROF. DR. CHRISTINE LEGNER, Competence Center Corporate Data Quality, Universität Lausanne:

*Data Governance als Voraussetzung für Datenqualität – Kann die Forschung von Unternehmen lernen?*

- PROF. DR. THOMAS GANSLANDT, Heinrich-Lanz-Center for Digital Health, Mannheim University Medicine/Universität Heidelberg:

*Daten-Qualität, -Provenance und -Governance: Kollaborative Ansätze in der Medizininformatik-Initiative des BMBF*

- DR. NIKLAS BLOMBERG, ELIXIR – verteilte Infrastruktur für lebenswissenschaftliche Informationen:

*Standards, Services and Scientific Communities: Foundations for a Sustained Federated Data Infrastructure*

**12:30 Uhr**    **Gemeinsames Mittagessen**

**13:00 Uhr**

**Podiumsdiskussion: Forschung – Daten – Qualität: Künftige Herausforderungen und Lösungsansätze**

Moderation:

- DR. JAN-MARTIN WIARDA, Journalist für Bildung und Wissenschaft

Podiumsteilnehmer:

- PROF. DR. LARS BERNARD, Dekan der Fakultät für Umweltwissenschaften der Technischen Universität Dresden, Mitglied des RfII
- PROF. DR. MARTINA BROCKMEIER, Vorsitzende des Wissenschaftsrates (WR) 2017 – 2020
- PROF. DR. FRAUKE KREUTER, Statistik und Methoden, Universität Mannheim und Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit (IAB)
- PROF. DR. KLAUS POHL, Direktor paluno – The-Ruhr-Institute for Software Technology, Universität Duisburg-Essen
- KATRIN STUMP, M.A., Direktorin der Universitätsbibliothek Braunschweig, Mitglied des RfII

**14:30 Uhr**    **Zusammenfassung der Konferenzergebnisse**

PROF. DR. PETRA GEHRING, Vorsitzende des RfII

**14:45 Uhr**    **Ende der Konferenz**