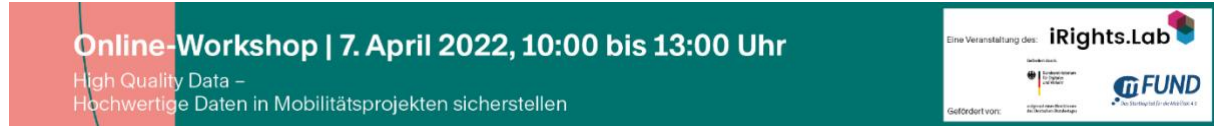


Emmett

Veranstaltungsdokumentation:



Beschreibung

Der Umgang mit Daten ist oftmals der gemeinsame Nenner innovativer Vorhaben für einen effizienten und nachhaltig gestalteten Verkehr auf der Straße, in der Luft, auf dem Wasser und auf der Schiene. Die Forschungs- und Entwicklungsteams hinter diesen Vorhaben stehen vor der Aufgabe, große Mengen an Daten zu erheben, zu analysieren und zu verwerten. Eine wiederkehrende Herausforderung dabei ist, sicherzustellen, dass die Daten bestimmten Qualitätsanforderungen entsprechen – insbesondere für Forschungsprojekte mit geringem Budget. Denn um die Qualität von Daten hinlänglich bewerten zu können, sind unter anderem rechtliche sowie technische Kenntnisse notwendig. Der Zugang zu dieser Expertise kann aufwendig und kostenintensiv sein.

Die Methoden und Instrumente, mit denen die Datenqualität gemessen werden kann, können unterschiedlich sein. Häufig hängt die Auswahl davon ab, aus welcher Perspektive die Datenqualität beurteilt wird. Zum Beispiel differieren in den Rechts- oder Informationswissenschaften die Definitionen von Datenqualität und setzen verschiedene Qualitätskriterien voraus.

Bei dem Workshop „High Quality Data – Hochwertige Daten in Mobilitätsprojekten sicherstellen“ konnten sich Vertreter*innen verschiedener Fachrichtungen interdisziplinär austauschen. In zwei Impulsvorträgen vermittelten die Referent*innen grundlegende Informationen zur Datenqualität im gesamten Prozess des Datenmanagements und stellten rechtliche und technische Lösungen für das Qualitätsmanagement vor.

Die Teilnehmer*innen diskutierten über Herausforderungen und Bedarfe in ihren eigenen Mobilitätsprojekten und tauschten sich über praktische Hinweise und nützliche Tools aus. Am Ende der Veranstaltung näherten sich die Teilnehmer*innen einer Definition von Datenqualität an – aus den Perspektiven Künstliche Intelligenz (KI) und Trainingsdaten, Infrastrukturdaten sowie Daten der öffentlichen Verwaltung.

Programm

TEIL 1 Begrüßung und Impulse I

- 10:00 Uhr** **Begrüßung und Einführung**
Mariel Sousa und Anne Lammers, Move Mobility, iRights.Lab
- 10:05 Uhr** **Begrüßung aus dem BMDV**
Rudolf Boll, Referat DP 24, BMDV
- 10:10 Uhr** **Kennenlernen**
- 10:15 Uhr** **Impulsvortrag: Die Rolle der Datenqualität im Data-Governance-Prozess**, Dr. Nikolai Horn, iRights.Lab

- 10:35 Uhr** Fragen, Antworten, Diskussion
- 10:40 Uhr** Impulsvortrag: Datenqualität am Beispiel des mFUND-Projekts SAUBER, Prof. Dr. Beatrix Weber, Hochschule für Angewandte Wissenschaften Hof
- 10:50 Uhr** Fragen, Antworten, Diskussion
- 11:00 Uhr** Pause
- TEIL 2 Impulse II und Gruppenarbeit
- 11:10 Uhr** Einführung in die Gruppenarbeit
- 11:20 Uhr** Gruppenarbeit in drei Breakout-Sessions
- Gruppe 1, Trainingsdaten und KI:** Moderation Move Mobility, inhaltliche Begleitung mFUND-Projekt [SAUBER](#)/ [KLIPS](#)
Hendrik Herold und Juliane Wright, Leibniz-Institut für ökologische Raumentwicklung
- Gruppe 2, Daten der öffentlichen Verwaltung:** Moderation Move Mobility, inhaltliche Begleitung mFUND-Projekt [AktMel](#)
Prof. Dr. Richard Goebel, Institut für Informationssysteme, Hochschule Hof
- Gruppe 3, Infrastrukturdaten:** Moderation Move Mobility, inhaltliche Begleitung mFUND-Projekt [ENGINEER](#)
Dr.-Ing. Jörg Bödefeld, Bundesanstalt für Wasserbau
- 12:20 Uhr** Pause
- TEIL 3 Austausch, Diskussion und Netzwerken
- 12:30 Uhr** Ergebnispräsentation
- 12:55 Uhr** Zusammenfassung und Ausblick
- 13:00 Uhr** Ende der Veranstaltung



Abbildung 1: In diesem Workshop sprachen 55 Teilnehmende über Datenmanagement und ihre praktischen Erfahrungen zur Sicherstellung einer hohen Datenqualität.

Die Referent*innen



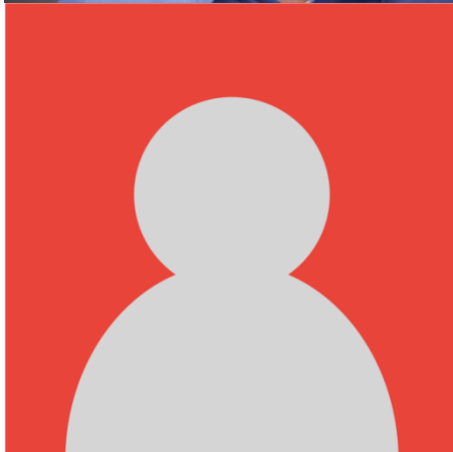
Dr. Nikolai Horn hat nach seinem Studium am Institut für Öffentliches Recht der Universität Bonn gearbeitet und wurde interdisziplinär am Lehrstuhl des Bundesverfassungsrichters Dr. Udo Di Fabio promoviert. Er beschäftigt sich seit Jahren neben- und hauptberuflich mit den Themen Digitalethik und Datenschutz. Nikolai Horn koordiniert beim iRights.Lab das Projekt Data Governance.



Prof. Dr. Beatrix Weber arbeitet mit ihrer Forschungsgruppe „Recht in Nachhaltigkeit, Compliance und IT“ am Institut für Informationssysteme der Hochschule für Angewandte Wissenschaften Hof. Ihre Mission dabei ist: „Innovative Technologien rechtlich möglich machen“. Recht und KI, Data Governance und Compliance sind ihre Schwerpunkte. Sie nutzt ihre Expertise unter anderem in der Arbeit an den mFUND-Projekten SAUBER und KLIPS.



Henrik Herold ist Projektleiter im Forschungsbereich „Raumbezogene Information und Modellierung“ am Leibniz-Institut für ökologische Raumentwicklung. Er forscht auf dem Gebiet der Geoinformatik unter anderem zu Methoden der Gewinnung und Analyse raumzeitlicher Informationen.



Juliane Wright studierte Raumplanung an der Technischen Universität Dortmund und arbeitete dort an der Erstellung großflächiger Klimawirkungsanalysen. Im Rahmen ihrer derzeitigen Tätigkeit am Leibniz-Institut für ökologische Raumentwicklung befasst sie sich mit der Planungspraxis und technischen Entwicklung neuer KI-Verfahren für die Hitzeanpassung.



Prof. Dr. Richard Göbel studierte Informatik an der TH Darmstadt. Er promovierte an der Universität Kaiserslautern im Bereich der Künstlichen Intelligenz. Seit 1997 ist Richard Göbel Professor an der Hochschule für Angewandte Wissenschaften Hof. Am Institut für Informationssysteme (iisys) leitet er die Forschungsgruppe „Multimediale Informationssysteme“.



Dr. Jörg Bödenfeld studierte an der Universität Wuppertal Bauingenieurwesen. Er promovierte in Leipzig im Bereich des Stahlbetonbaus. Seit 1994 ist er bei der Bundesanstalt für Wasserbau tätig. Dort leitet er das Referat „Infrastrukturmanagement“.

Zusammenfassung

Zusammenfassung des Impulsvortrags: Datenqualität und Data Governance

Um die Weiterverarbeitung von Daten zu ermöglichen, ist eine hohe Datenqualität von zentraler Bedeutung. Gängige Definitionen beziehen die Qualität von Daten auf den für sie vorgesehenen Gebrauch beziehungsweise ihre Eignung für ein bestimmtes Ziel im operativen Geschäft. Beispiele für Kriterien zur Bewertung der Qualität sind die Korrektheit, Relevanz und Verlässlichkeit von Daten.

Im operativen Geschäft ist es allerdings häufig eine Herausforderung, den Überblick zu behalten und Prüfsysteme für Daten einzuführen. Oftmals fehlen die notwendigen Ressourcen für ein geeignetes Datenmanagement, insbesondere in Forschungs- und Entwicklungsprojekten mit geringem Budget. Transparentes Wissen und Methoden zur Selbsthilfe können hilfreiche Werkzeuge liefern, um Datenqualität zu bewerten und zu verbessern. Dr. Nikolai Horn war an der Entwicklung des [Self-Data-Governance-Frameworks](#) beteiligt, das Projekten eine Anleitung für ein eigenes Datenmanagement bietet. Dabei stehen der Inhalt, die Nutzbarkeit und das Management der Daten im Fokus.

Prof. Dr. Beatrix Weber nutzt das Framework unter anderem im mFUND-Projekt [SAUBER](#). Die übersichtliche Tabelle bietet die Möglichkeit, die Datenqualität einzuschätzen sowie greifbar und vergleichbar zu machen. In der Praxis lässt sich das Framework auf verschiedene Forschungs- oder Entwicklungsprojekte anwenden.

Die beiden Expert*innen waren sich einig, dass der Data Act, der aktuell von der EU-Kommission ausgearbeitet und diskutiert wird, die Arbeit mit Daten grundlegend verändern könnte. Er würde die Interoperabilität fördern, wodurch die Datenqualität als Voraussetzung für die Weiterverwendung von Daten an Bedeutung gewinnen würde.

Zusammenfassung der Gruppenarbeiten

Nach der inhaltlichen Einführung arbeiteten die Teilnehmenden in Gruppen an der Rolle von Datenqualität und Data Governance in unterschiedlichen Arbeitsbereichen. So ging es um die Anforderungen, die an die Datenqualität bei der Arbeit mit Trainingsdaten und KI, mit Infrastrukturdaten sowie in der öffentlichen Verwaltung gestellt werden. Außerdem tauschten sich die Teilnehmenden zu Herausforderungen und Bedarfen bei ihrer Arbeit mit Daten aus.

Die Ergebnisse der Gespräche hielten sie in einem Miro-Board fest:

<https://miro.com/app/board/uXjVOA2MDUI=/>.

Im unteren Teil des Miro-Boards fassten die Gruppen für ihre jeweiligen Bereiche zusammen, wie sie Datenqualität definieren würden und welche wesentlichen Bedarfe und Herausforderungen ihnen in ihrem Bereich begegnen.

Im Folgenden werden die Ergebnisse der Gruppenarbeiten zusammengefasst.

Gruppe 1) Trainingsdaten und KI

Bei der Arbeit mit Trainingsdaten und Künstlicher Intelligenz gibt es einen besonders hohen Aufwand bei der Datenaufbereitung. Damit die Daten mit einer KI verarbeitet werden können, müssen sie homogenisiert werden. Außerdem ist die Genauigkeit der Daten wichtig. Sollen aktuelle Daten stetig verarbeitet werden, müssen geeignete Datenströme für Echtzeit-Anwendungen geschaffen werden. Bei diesen Herausforderungen hilft es, vorab Standards für die Qualität definiert zu haben. Nur so könnten die Ergebnisse auch nach dem Projekt weiterverwendet werden.

Für die Teilnehmenden, die mit Trainingsdaten und KI arbeiten, ist Datenqualität essenziell. Kriterien für eine hohe Datenqualität sind die Vertrauenswürdigkeit, die zeitliche und räumliche Genauigkeit sowie die Vollständigkeit der Daten.

Gruppe 2) Infrastrukturdaten

Um Daten, die die Infrastruktur betreffen, mit einer hohen Qualität verarbeiten zu können, werden Metriken zur Bestimmung der Datenqualität benötigt. Bei verschiedenen Akteur*innen ist es eine Herausforderung, dass alle Beteiligten über das gleiche Wissen zur Datenverarbeitung verfügen und Datenqualität als gewolltes Ziel auf allen Ebenen feststeht. Auch in diesem Bereich helfen Datenstandards dabei, eine hohe Datenqualität flächendeckend zu erreichen.

Aus Sicht derer, die mit der Infrastruktur arbeiten, ist Datenqualität der entscheidende Faktor für das Ergebnis der Datennutzung. Da im Bereich der Infrastruktur immer verschiedene Akteur*innengruppen zusammenarbeiten, müssen die Zuständigkeiten klar verteilt sein. Für die Arbeit mit denselben Daten in verschiedenen Stellen müssen außerdem Gesetze und technische Normen den Standard festlegen. Die IT-Systeme sollten so offen sein, dass Daten schnell eingebunden werden können und so Interoperabilität ermöglicht wird. Die Aktualität ist zudem eine wichtige Voraussetzung, um die Daten nutzen zu können.

Gruppe 3) Daten der öffentlichen Verwaltung

Für eine hohe Qualität der Daten, die in der öffentlichen Verwaltung genutzt werden, ist die Aktualität der Daten eine Herausforderung. Außerdem müssen die Daten validiert und auf

die Einhaltung von Metadatenstandards – also die Art und Weise, wie die Daten in Datenbanken abgelegt werden – geprüft werden.

Für die öffentliche Verwaltung sind bei der Bewertung der Datenqualität juristische Kriterien im Allgemeinen wichtig. Zunächst müssen die Arten von Daten unterschieden werden. Deren Dokumentation sollte immer parallel stattfinden. Um Daten weiterverwenden zu können, müssen die Nutzungsrechte gewährleistet sein, ggf. muss die Freigabe durch den Datenowner – also die jeweilige öffentliche oder private datenhaltende Stelle – erfolgen. Anschließend sollte grundsätzlich die Eignung der Daten für den Anwendungsfall geprüft werden. So ist Datenqualität für die öffentliche Verwaltung nur mit Blick auf die spätere Anwendung zu definieren und nur unter Einhaltung der gesetzlichen Vorgaben herzustellen.

Weiterführende Links und Literatur

- Was bedeutet „Data Governance“?: <https://emmett.io/article/was-bedeutet-data-governance>
- Self-Data-Governance-Framework: <https://emmett.io/article/self-data-governance-framework>
- Projekt Data Governance am iRights.Lab: <https://www.irights-lab.de/projekt/data-governance>
- mFUND-Projekt SAUBER: <https://www.bmdv.bund.de/SharedDocs/DE/Artikel/DG/mfund-projekte/sauber.html>
- mFUND-Projekt KLIPS: <https://www.bmdv.bund.de/SharedDocs/DE/Artikel/DG/mfund-projekte/klips.html>
- mFUND-Projekt AktMel: <https://www.bmdv.bund.de/SharedDocs/DE/Artikel/DG/mfund-projekte/akt-mel.html>
- mFUND-Projekt ENIGINEER: <https://bmdv.bund.de/SharedDocs/DE/Artikel/DG/mfund-projekte/engineer.html>
- Vorschlag zum Data Act der Europäischen Kommission: https://ec.europa.eu/commission/presscorner/detail/de/ip_22_1113

Unsere Online-Workshops

Die Online-Workshops sind auf Interaktion unter den Teilnehmenden ausgerichtet. Mitarbeitende aus mFUND-Projekten tauschen sich über die Arbeit in ihren Vorhaben aus, lernen sich kennen und vernetzen sich. Im Verlauf der etwa zwei- bis dreistündigen Online-Workshops werden praxisrelevante Informationen für die Projektarbeit vermittelt, stellen mFUND-Beteiligte Best-Practices und Erfahrungsberichte vor und externe Expert*innen geben fachspezifischen Input. Neben Kurzpräsentationen und Werkstattberichten finden Gruppenarbeiten in Breakout-Sessions beziehungsweise Diskussionen und interaktiver Austausch statt. Informationen und Handreichungen für die Projektarbeit werden während der Veranstaltung zur Verfügung gestellt sowie im Nachgang auf emmett.io veröffentlicht, zum Beispiel die Emmett-Wirkungsmatrix und der Leitfaden zur Wirkungsevaluation für datengetriebene Mobilitätsprojekte (PDF). Die Workshops wenden sich an mFUND-Projektmitarbeiter*innen und die Teilnehmendenzahl ist auf circa 40 begrenzt.

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages



Über Emmett und Kontakt

Emmett ist eine offene Kommunikations- und Vernetzungsplattform für datengetriebene Mobilitätsprojekte, initiiert und umgesetzt vom unabhängigen Thinktank iRights.Lab. Die Plattform bietet eine Übersicht und einen Einblick in die Projekte der Forschungsinitiative mFUND (Modernitätsfonds) des Bundesministeriums für Digitales und Verkehr (BMDV). Emmett dient der Vernetzung und dem Austausch von Wirtschaft, Wissenschaft und Politik sowie allen Projektteilnehmer*innen und interessierten Bürger*innen. Die Plattform gibt einen branchenübergreifenden Einblick in die aktuelle Forschung und Entwicklung innovativer Mobilität in Deutschland.

Neuigkeiten zur datengetriebenen Mobilität und der mFUND-Begleitforschung:

Webseite: www.emmett.io

Twitter: <https://twitter.com/emmettmobility>

LinkedIn: <https://www.linkedin.com/company/emmettmobility/>

Für Fragen zu den Fachaustauschen wenden Sie sich gerne an das Emmett-Team:

momo@irights-lab.de

Informationen zum mFUND-Frauen Netzwerk *Women in Datadriven Mobility* (WDM):

<https://emmett.io/article/das-potenzial-geschlechterspezifischer-daten-fuer-mobilitaetsprojekte>