

Angelika Schulz, Elmar Brockfeld, René Kelpin und Antje Parnitzke
Deutsches Zentrum für Luft- und Raumfahrt e. V. (DLR)
- Institut für Verkehrsforschung (IVF) -
Rutherfordstraße 2, 12489 Berlin

Dateninformationsangebot der Clearingstelle für Verkehrsdaten und Verkehrsmodelle des DLR

In der Verkehrsforschung und -planung besteht erheblicher Bedarf an detaillierten Informationen über das Verkehrsgeschehen. Entsprechend zahlreich und vielfältig sind empirische Studien, Messkampagnen, Modellansätze und Simulationsprogramme. Im Hinblick auf die Nutzung bereits vorhandener Informationen erweisen sich insbesondere heterogene Datenformen, unvollständig beschriebene Datensätze oder fehlende Hinweise auf die Qualität der jeweiligen Daten als problematisch. Ähnliches gilt für den Bereich der Verkehrsmodelle¹. Zudem sind zahlreiche Arbeiten nur einem vergleichsweise begrenzten Nutzerkreis bekannt und zugänglich, so dass wertvolle Ressourcen häufig ungenutzt bleiben.

Die im DLR-Institut für Verkehrsforschung (IVF) angesiedelte Clearingstelle für Verkehrsdaten und -modelle² möchte mit dem Aufbau einer zentralen Datenbasis diese Informationslücke schließen.

Thematischer Fokus: Verkehr und Mobilität

Der Bereich der Verkehrsdaten ist durch eine große Vielfalt gekennzeichnet. Neben Daten zum Verkehrsfluss, die in Messkampagnen erhoben werden, gibt es eine Fülle von Untersuchungen, deren Ziel beispielsweise die Ermittlung verkehrsrelevanter menschlicher Verhaltensaspekte oder individueller Entscheidungsgrundlagen ist. Solche Daten werden häufig durch Befragungen erhoben und statistisch ausgewertet. Als Beispiel seien hier die Studien "KONTIV"/ "Mobilität in Deutschland", das "Deutsche Mobilitätspanel" (MOP) oder das "System repräsentativer Verkehrsbefragungen" (SrV) genannt.

Das Datenangebot der Clearingstelle konzentriert sich zunächst auf Sondererhebungen, die im Auftrag der Bundesministerien für Verkehr, Bau- und Wohnungswesen (BMVBW) sowie Bildung und Forschung (BMBF) durchgeführt werden. Das Angebot soll nach und nach erweitert werden. Im Hinblick auf Daten der amtlichen oder halb-amtlichen Statistik wird auf die bestehenden Informationsportale der entsprechenden Stellen verwiesen.

Metadaten und Daten

Die erfolgreiche Nutzung einmal erhobener Daten hängt – insbesondere im Hinblick auf Sekundäranalysen – wesentlich von einer sorgfältigen Dokumentation sowie geeigneten Publikationsformen ab. Für die Auswahl der geeigneten empirischen Datenbasis sowie deren fachgerechte Analyse sind detaillierte Informationen zum Erhebungskontext, etwa zur Methodik, unabdingbare Voraussetzung. Eine vorrangige Aufgabe der Clearingstelle ist deshalb die Bereitstellung

¹ Auf den Bereich der Verkehrsmodelle wird im Folgenden nicht eingegangen.

² Das Institut für Verkehrsforschung (IVF) des Deutschen Zentrums für Luft- und Raumfahrt (DLR) betreibt die Clearingstelle im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF).

ausführlicher Metadaten, die sowohl über die jeweiligen Datenbestände als auch über Nutzungsmöglichkeiten und Bezugsquellen Auskunft geben.

Metadaten sind Daten über Daten. Der Definition der ISO/IEC 11179-1 entsprechend handelt es sich um jene dokumentierten Informationen, die das jeweils beschriebene Objekt, z. B. den Datenbestand einer empirischen Erhebung verstehbar und somit interpretierbar machen. Zu den häufig verwendeten Metadatenelementen gehören neben bibliographischen Angaben wie Titel, Autor, Auftraggeber oder Jahr der Veröffentlichung v. a. eine Zusammenfassung sowie Schlagwörter. Werden Metadaten systematisch erfasst und strukturiert abgelegt, ermöglichen sie mittels geeigneter Suchmaschinen einerseits komfortable Recherchen auch in umfangreichen Metadatenbeständen. Eine einheitliche, standardisierte Beschreibung erleichtert zudem den einfachen, direkten Vergleich verschiedener Objekte.

Die Dokumentation der Datenbestände, die über die Clearingstelle zur Verfügung gestellt werden, orientiert sich an einem Metadatenstandard, der von der internationalen *Data Documentation Initiative* für die Dokumentation empirischer sozialwissenschaftlicher Studien entwickelt wurde (DDI 2002). Basis ist die *Document Type Definition "codebook.dtd"*, in der eine Vielzahl möglicher Metadatenelemente definiert ist. Diese Zusammenstellung umfasst weit mehr Elemente als die üblichen Quellenangaben. Sie lassen sich in fünf Gruppen zusammenfassen:

DDI-Codebook-Abschnitt	Inhalt
Document Description	beschreibt das Metadatendokument
Study Description	beschreibt die zugrunde liegende Studie
File Description	beschreibt die zugrunde liegende Datendatei, z.B. eine SPSS-Datei (*.sav)
Data/ Variable Description	beschreibt die Struktur der Daten (Variablen, Label usw.)
Study-Related Material	enthält Hinweise auf ergänzendes Material

Die inhaltliche, methodische und formale Beschreibung ist in den beiden Abschnitten *Study Description* und *Data/ Variable Description* enthalten. Die Studienbeschreibung umfasst u. a. die herkömmlichen Quellenangaben und gibt darüber hinaus Auskunft über zeitliche und räumliche Abdeckung, die Grundgesamtheit, das Studiendesign, die Art der Datenerhebung, den realisierten Rücklauf, durchgeführte Maßnahmen der Datenweiterbearbeitung, angewendete Gewichtungsverfahren sowie die Verfügbarkeit der Daten und Nutzungsbedingungen. Die Daten- und Variablenbeschreibung beinhaltet neben den gebräuchlichen Angaben eines Codeplans beispielsweise auch die zugrunde liegenden Quellfragen oder Intervieweranweisungen. Die direkt auf die jeweilige Studie bezogenen Metainformationen werden gegebenenfalls durch ergänzende Metadatendokumente und weiterführende Verweise (Hyperlinks) komplettiert. Dies können gescannte Fragebögen oder Literaturangaben sein. Sofern die Autoren oder Dateneigentümer ein eigenes Informationsportal anbieten, wird auf dieses ebenfalls verwiesen.

Alle Metainformationen werden i. d. R. von Mitarbeitern der Clearingstelle zusammengestellt. Als Grundlage der Dokumentation dienen von den Dateneigentümern veröffentlichte oder bereitgestellte

Materialien wie beispielsweise Studienabschluss- und Methodenberichte, aufbereitete Rohdatenbestände oder Codepläne.

In Ergänzung zur Metadatendokumentation werden einige Datenbestände über den NESSTAR-Server³ der Clearingstelle zugänglich gemacht, die einen ersten Einblick in den jeweiligen Datensatz sowie elementare Analysen erlaubt (z. B. ungewichtete Häufigkeitsauszählungen, Kreuztabellen). NESSTAR unterstützt ebenfalls den Metadatenstandard der DDI.

Zugang zu Metadaten und Datensätzen

Die Kataloge, Linksammlungen und Metainformationen der Clearingstelle sind frei zugänglich und ohne Einschränkung nutzbar. Datenbestände hingegen sind in der Regel zwar für wissenschaftliche, nichtkommerzielle Nutzung in Lehre und Forschung freigegeben, ihre Nutzung ist jedoch häufig an besondere Auflagen gebunden und nur mit Zustimmung des Dateneigentümers möglich.

Für den direkten Zugriff auf Daten mit besonderen Zugangsbedingungen ist deshalb eine Anmeldung bei der Clearingstelle erforderlich. Informationen zu den aktuellen Zugangsbedingungen können den jeweiligen Metadaten entnommen werden. Sind Zugang und Nutzung an die Zustimmung des Eigentümers gebunden, wird diese in jedem Fall vor einer Freischaltung von der Clearingstelle eingeholt.

Zielgruppe

Die Clearingstelle richtet sich sowohl an Einrichtungen, die verkehrsrelevante Daten erheben als auch an jene, die diese Forschungsergebnisse nutzen möchten. Sie versteht sich in erster Linie als Mittler zwischen Anbietern von Daten einerseits und deren Nutzern andererseits. Anbieter erhalten ein zusätzliches Forum für die Publikation ihrer Arbeiten. Nutzern hingegen wird eine Anlaufstelle angeboten, die sie bei der Suche nach Informationen und Ansprechpartnern unterstützt.

Insbesondere im Hinblick auf die Bereitstellung von Daten ist die Clearingstelle nicht zuletzt auf die Bereitschaft der Datenproduzenten angewiesen, ihre jeweiligen Datenbestände – versehen mit der erforderlichen Dokumentation – Dritten zur Verfügung zu stellen.

Literatur

DDI (Data Documentation Initiative) (2002): **Codebook DTD. Version 1.2.2 vom 28.08.2002.**

<http://www.icpsr.umich.edu/DDI/CODEBOOK/codebook1-2-2.html>

ISO /IEC 11179-1:1999(E): **Information technology – Specification and standardization of data elements – Part 1: Framework for the specification standardization of data elements.**

http://isotc.iso.ch/livelink/livelink/fetch/2000/2489/Ittf_Home/PubliclyAvailableStandards.htm??Redirect=1

NESSTAR (Network Social Science Tools and Resources): <http://www.nesstar.com>

³ NESSTAR: Network Social Science Tools and Resources