

# Erfahrungen bei der Umsetzung eines Datenmanagementplans für räumliche Daten des Radverkehrs

## *Experiences in the Implementation of a Data Management Plan for Spatial Data of Bicycle Traffic*

Sven Leitinger<sup>1</sup>, Andreas Wagner<sup>1</sup>, Wolfgang Kremser<sup>1</sup>

<sup>1</sup>Salzburg Research Forschungsgesellschaft mbH · sven.leitinger@salzburgresearch.at

**Zusammenfassung:** Im Projekt *Bicycle Observatory* werden eine Vielzahl an unterschiedlichen räumlichen Datensätzen des Radverkehrs in Beziehung gesetzt und die Ergebnisse sollen, wo aus lizenzrechtlicher Sicht möglich, als Open Data der wissenschaftlichen Community bereitgestellt werden. Um diese Anforderungen über den kompletten Projektzyklus zu unterstützen hat sich in den letzten Jahren das Instrument des Datenmanagementplans auch in Forschungsprojekten etabliert. Zur Unterstützung dieser Aufgaben wurde auf Basis der Kollaborationsplattform *Confluence* ein leicht pflegbarer Datenmanagementplan aufgesetzt.

**Schlüsselwörter:** Mobilität, Radverkehr, Räumliche Daten, Datenmanagement

**Abstract:** *In the project, Bicycle Observatory a multitude of different spatial data sets of bicycle traffic will be correlated and the results will be made available to the scientific community as open data, where possible from a licensing perspective. In order to support these requirements over the complete project cycle, the instrument of the data management plan has also been established in research projects in recent years. To support these tasks, an easy to maintain data management plan has been set up based on the collaboration platform Confluence.*

**Keywords:** *Mobility, cycling, spatial data, data management*

## 1 Einleitung

Das wachsende Bewusstsein für Privatsphäre in der Zeit von Big Data schafft neue Rahmenbedingungen, auch für Forschungsprojekte. Eine steigende Anzahl von Fördergebern fordert mittlerweile einen Datenmanagementplan (DMP). Zum einen soll der DMP Personen vor ausufernder Datensammlung schützen, andererseits dient er zur Sicherstellung der Datenqualität und der Replizierbarkeit von Forschungsergebnissen.

Mit dieser Problematik wurden wir im Rahmen des Forschungsprojektes *Bicycle Observatory* konfrontiert. Darin wird erstmals versucht das von Miller (2017) vorgestellte Konzept eines *Geographic Information Observatory* am Beispiel des Radverkehrs umzusetzen (siehe auch Loidl et al., 2020, Brocza et al., 2020).

Das *Bicycle Observatory* hat das Ziel verschiedene räumliche Datensätze unterschiedlicher Herkunft (verschiedene Dateneigentümer und Nutzungsverträge) und Thematik (z. B. Rad-Infrastruktur, Sensordaten, sozialwissenschaftliche Erhebungen etc.) zum Radverkehr in Beziehung zu setzen und in einem mehrdimensionalen Lagebild darzustellen. Ein weiteres Ziel ist es, die im Projekt aggregierten Daten als Open Data zur Verfügung zu stellen.

Diese Ziele erfordern ein durchgängiges Management aller untersuchter und verwendeter Datensätze. Es erscheint sinnvoll alle Datensätze in einem DMP abzubilden. Der DMP beschreibt welche Daten erhoben bzw. integriert werden, wie die Daten gespeichert, gesichert und gepflegt werden. Darüber hinaus wird die Art der Datendokumentation (Metadaten), die Lizenzbedingungen (die auch den Zugriff Dritter auf die Daten regeln) und die Verantwortlichkeit für die Daten festgehalten.

Der vorliegende Artikel beschreibt die methodischen Überlegungen zur Umsetzung eines DSGVO-konformen DMP für räumliche Daten. Auf Basis dieser generischen Überlegungen wird anhand des Projektes Bicycle Observatory ein Ansatz für eine einfache Umsetzung und Pflege des DMP gezeigt. Abschließend werden die Erkenntnisse aus der bisherigen operativen Befüllung und Pflege des DMP zusammengefasst.

## 2 Grundlagen und Anforderungen an das Datenmanagement

Der Begriff Datenmanagement umfasst die Planung, Durchführung und Überwachung von Prozessen und Richtlinien bei der Verwertung von Daten. Datenmanagement ist im gesamten Daten-Lebenszyklus notwendig: Aufzeichnung, Speicherung, Wartung, Verwendung und eventuelle Löschung der Daten. Daher erfordert Datenmanagement sowohl technische als auch organisatorische Ressourcen (DAMA International, 2017).

Im Gegensatz dazu ist der Begriff DMP nicht eindeutig definiert, weder in bestehender Literatur noch vom Fördergeber selbst. Format sowie Inhalt des DMP sind somit frei wählbar. Ohne klare Vorgaben orientieren sich viele Forschungsprojekte – wie auch das Projekt *Bicycle Observatory* – bei der Erstellung ihres DMP an der Datenschutz-Grundverordnung (DSGVO) (EU, 2016). Unter diesem Blickwinkel bewerten wir einen DMP genau dann als ausreichend, wenn er die Konformität mit der DSGVO sicherstellt.

### 2.1 Minimale Anforderungen an den Datenmanagementplan

Die DSGVO regelt ausschließlich den Umgang mit personenbezogenen Daten. Art. 4 bezeichnet Daten als personenbezogen, sobald mit ihnen eine Zuordnung zu (u. a.) Standortdaten hergestellt werden kann. Somit ist dieser Aspekt für Daten mit räumlichen Bezug durchaus relevant, auch wenn augenscheinlich kein direkter Personenbezug (Name etc.) gegeben ist.

Laut Dove (2018) erhält die aufgezeichnete Person („*betreffene Person*“) eine Reihe von Rechten, denen die Person mit Verantwortung über die Datenverarbeitung („*Verantwortlicher*“) nachzukommen hat: (I) Recht auf Transparenz (Art. 5), (II) Auskunftsrecht (Art. 15) und (III) Recht auf Löschung (Art. 16).

Aus den vorher genannten Pflichten für den Verantwortlichen ergibt sich nun der minimal notwendige Inhalt des DMP (siehe Tabelle 1).

### 2.2 Metadatenmanagement und der Bezug auf existierende Metadaten-Schemata

Neben der reinen Erfüllung der Anforderungen der DSGVO bietet ein DMP eine zentrale Stelle für ein einheitliches Management der Metadaten von verwendeten Daten eines Forschungsprojekts. In Österreich wurde dazu von der Cooperation OGD Österreich, Arbeitsgruppe Metadaten (2017) das Schema „*OGD Metadaten 2.4*“ ausgearbeitet. Dabei werden

auch die Anforderungen von *INSPIRE* (INSPIRE, 2020), der *Public-Sector-Information-Richtlinie* (digitales.oesterreich, 2020) und des *Data Catalogue Vocabulary Application Profile* (DCAT-AP, 2020) berücksichtigt. *OGD Metadaten 2.4* wird in Österreich als de-Facto Standard für die Verarbeitung von Metadaten angesehen.

*OGD Metadaten 2.4* definiert eine Liste von Pflichtfeldern und optionalen Feldern zur semantischen Beschreibung von Daten. Die Verwendung dieses bestehenden Vokabulars verringert den Schritt hin zu einer Veröffentlichung der Datensätze, da Datenaustauschplattformen bereits damit umgehen können.

Durch das Vorliegen von *OGD Metadaten 2.4* erscheint es sinnvoll in den DMP auch noch die Anforderungen des Metadatenmanagements aufzunehmen. So bildet *OGD Metadaten 2.4* bereits die unter Kapitel 2.1 definierten minimalen Anforderungen des DMP für die DSGVO ab. Zusätzlich werden auch noch alle weiteren definierten Pflichtfelder des Metadatenkerns und soweit sinnvoll auch ausgewählte zusätzliche optionale Felder in den DMP aufgenommen (siehe Tabelle 1).

### 2.3 Weitere Felder im Datenmanagementplan

Neben *OGD Metadaten 2.4* können bei der Umsetzung eines DMP noch weitere Quellen (z. B. EC 2020, ICPSR 2012) für die Anforderungen hinsichtlich räumlicher Datenverarbeitung berücksichtigt werden. Dabei hat sich herausgestellt, dass die Checkliste zur Erstellung von Datenmanagementplänen vom Verbund Forschungsdaten Bildung (VerbunFDB, 2015) eine hilfreiche und ergänzende Grundlage ist. Diese werden in weiterer Folge als projektspezifische Felder im DMP abgebildet (siehe Kapitel 3.1).

Vor allem in Bezug auf personenbezogene Daten ist es hier sinnvoll weitere Punkte im DMP zu dokumentieren. Werden im Projekt personenbezogene Daten verwendet, so sollen die Maßnahmen zum verantwortlichen Umgang dieser Daten dokumentiert werden. Kann der Personenbezug bei der geplanten Datenverarbeitung vermieden werden sollte zur Dokumentation, soweit bekannt, die Anonymisierung dieser Daten festgehalten werden.

## 3 Umsetzung eines Datenmanagementplans im Projekt *Bicycle Observatory*

Im Projekt *Bicycle Observatory* werden eine Vielzahl von Datenschichten aus verschiedenen Datenquellen integriert, aufbereitet und weiterverarbeitet. Die Lizenzen der integrierten Datenschichten variieren von Open Data bis zu kommerziellen Lizenzen.

Während alle Datensätze einen räumlichen Bezug aufweisen (Loidl et al., 2020), wurde sichergestellt, dass zum aktuellen Projektzeitpunkt in keinen der Datensätze ein Personenbezug hergestellt werden kann. Somit unterliegen die Datensätze des *Bicycle Observatory* aus Sicht des Personenbezugs aktuell nicht der DSGVO.

Im Konsortium des Projekts *Bicycle Observatory* hat man sich dennoch darauf verständigt, einen, auf Basis des in Kapitel 2 entwickelten Grundschemas, DSGVO-konformen DMP für räumliche Daten umzusetzen. Neben dem Management der Metadaten, Lizenzbedingungen und Zuständigkeiten wurde es als wesentlicher Punkt erachtet, einen Weg des Datenmanage-

ments aufzusetzen, der auch für zukünftige Phasen des Projekts bzw. zukünftige Forschungsprojekte skaliert.

In den folgenden Abschnitten wird ein Einblick in die Entwicklung des DMP des *Bicycle Observatory*, sowie der für die Erstellung und Pflege notwendigen Prozesse gegeben.

**Tabelle 1:** Inhalte des DMP für räumliche Daten unter Berücksichtigung der DSGVO, *OGD Metadaten 2.4* und der Checkliste von Verbund Forschungsdaten Bildung

Anforderungen an DMP	Informationen	OGD Metadaten 2.4 Kurzname
Minimale Anforderungen (DSGVO)	Titel; Datenverantwortliche Stelle + Kontaktseite + E-Mailkontakt; Lizenz; Datenqualität/Herkunft; Attributbeschreibung; Zeitliche Ausdehnung Anfang + Ende; Datensatz, Dienst oder Dokument Link;	Title; maintainer + maintainer_email + maintainer_link; license; lineage_quality; attribute_description; begin_datetime + end_datetime; resource_url;
OGD Metadatenkern	Eindeutiger Identifikator; Datum des Datensatzes; Beschreibung; Kategorie; Schlagworte; Datensatz, Dienst oder Dokument Format; Veröffentlichende Stelle;	metadata_identifier; metadata_modified; description; categorization; keywords; resource_format; publisher;
OGD Optional	Weiterführende Metadaten + Link; Datensatz, Dienst oder Dokument Änderungsdatum; Geographische Abdeckung/Lage/Ausdehnung; Aktualisierungszyklus; Lizenz Link + Zitat; Veröffentlichende Stelle Kontaktseite + E-Mail; Link zu den Nutzungsbedingungen;	metadata_linkage_name + metadata_linkage; resource_lastmodified; geographic_toponym; geographic_bbox; update_frequency; license_citation + license_url; publisher_link + publisher_email; terms_url;
Projektspezifische Felder	Datenart; Maßnahmen bei personenbezogenen Daten / Anonymisierung; Verantwortlicher Projektpartner; Nutzen und Zweck	

### 3.1 Initialversion und erste Befüllung des DMP

Initial wurde eine Anpassung des generellen Grundschemas eines DMPs für das *Bicycle Observatory* durchgeführt. Konkret wurden notwendig bzw. sinnvolle optionale Felder des OGD Metadatenkerns identifiziert. Ferner wurde projektspezifische Felder festgelegt die in *OGD Metadaten 2.4* nicht abgebildet sind.

Mit dem angepassten Grundschema wurden eine Erstversion des DMP in Form eine Excel-Tabelle umgesetzt. Die Felder des DMPs wurden als Spalten modelliert, jeder Datensatz wird durch eine Zeile der Tabelle repräsentiert. In dieser Form wurde anschließend eine Befüllung mit den vorhandenen Datensätzen durchgeführt.

Bereits mit einem relativ geringen Datenbestand wurde klar, dass diese Umsetzung des DMP in einer Excel-Tabelle aus folgenden Gründen im weiteren Projektverlauf nur schwer integrierbar ist:

- Aufgrund der vielen Spalten und Einträge wurde die Tabelle schnell sperrig und unübersichtlich. Entsprechen aufwendig ist die operative Pflege.

- Die Tabelle und damit der ganze DMP kann immer nur von einer Person bearbeitet werden.
- Die aktuelle Version muss für Projektpartner explizit zur Verfügung gestellt werden (z. B. über eine gemeinsame Projektplattform).
- Beziehungen zwischen Datensätzen sind nur schwer abbildbar.

### 3.2 Onlinefähiger DMP auf Basis von Confluence

Aufgrund der festgestellten Schwachpunkte bei der Umsetzung des DMPs als Excel-Tabelle wurde im gegenständlichen Projekt beschlossen, die Strategie für die Umsetzung des DMP zu wechseln. Es wurde die Verwendung existierender Enterprise Tools für das Metadatenmanagement und die Abbildung des DMPs in Betracht gezogen (z. B. Dataspot<sup>1</sup>, Talend Data Catalogue<sup>2</sup>). Hier wurde festgestellt, dass die Verwendung eines externen Tools zu einer Reihe von neuen Fragestellungen führt. Diese reichen von der Evaluierung der notwendigen Funktionalität, der Einarbeitungszeit um das in Kapitel 2 vorgeschlagene Schema abzubilden bis hin zu einem DSGVO-konformen Betrieb der entsprechenden Tools.

Eine kostengünstigere Alternative ist die Nutzung von online DMP Assistenten (z. B. DMPOnline<sup>3</sup>, DMPTool<sup>4</sup>). Diese Webservices basieren auf Checklisten die von verschiedenen Open-Data-Initiativen, Forschungsorganisationen und Fördergebern formuliert wurden. Sie bieten Institutionen auch die Möglichkeit, Vorlagen und Guidelines hochzuladen. Allerdings sind im Moment keine Vorlagen aus Organisationen aus Österreich oder EU-Ländern hinterlegt. Eine stärkere Bewerbung und Nutzung heimischer Fördergeber würde die Nutzung dieser Tools attraktiver machen. Ein weiterer Nachteil ist der geschlossene Charakter dieser Systeme, der eine Automatisierung von DMP Tätigkeiten ausschließt.

Schlussendlich wurde in *Bicycle Observatory* die Umsetzung des DMP auf Basis der Kollaborationsplattform *Confluence*<sup>5</sup> als am zielführendsten erachtet. Dafür sprach eine Reihe (teils pragmatischer) Gründe: Da bei einem Projektpartner bereits eine DSGVO konforme Eigeninstallation von *Confluence* vorhanden war, konnten die technischen Aufwände für die Umsetzung des DMP minimiert werden. Die allgemeine Verwendung als Dokumentations-tool war ebenfalls bekannt. Zusätzlich ist *Confluence* mandantenfähig, das heißt es können auf verschiedenen Inhaltsebenen (Bereiche, Seiten und Unterseiten) Zugriffsrechte (lesen, bearbeiten) individuell oder auf Benutzergruppenebene vergeben werden. Die Exportfunktion generiert PDFs mit dem Inhalt einzelner Seiten oder Seitengruppen. Mit der eingebauten Suchfunktion können Einträge einzelner Datensätze wiedergefunden werden.

Spezielle DMP Funktionalität konnte mittels Vorlagen umgesetzt werden. Die Grundlage des *Confluence*-basierten DMP bildet eine Vorlage zur Erfassung der Datensätze (siehe Abbildung 1). Sie richtet sich nach den in Tabelle 1 definierten Feldern. Mit Hinblick auf die Zukunft bietet *Confluence* eine Plug-in-Architektur, die es grundsätzlich erlaubt, komplexere, DMP spezifische Aufgaben zu implementieren. Drittanwendungen könnten über die REST-Schnittstelle automatisch den DMP aktualisieren.

---

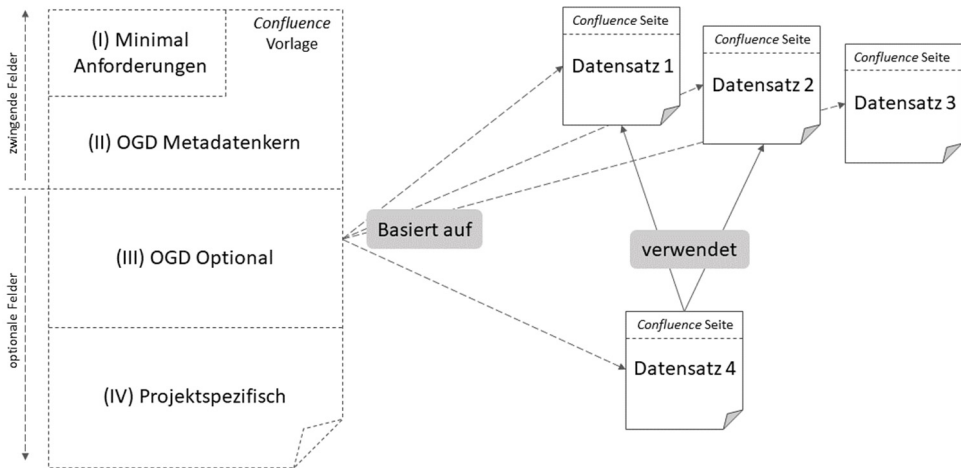
<sup>1</sup> <https://www.dataspot.at/software/>

<sup>2</sup> <https://de.talend.com/products/data-catalog/>

<sup>3</sup> <https://dmponline.dcc.ac.uk/>

<sup>4</sup> <https://dmptool.org/>

<sup>5</sup> <https://www.atlassian.com/de/software/confluence>



**Abb. 1:** Struktur des DMP in *Confluence* (eigene Abbildung)

### 3.3 Pflege des DMP im Projektkonsortium

Im Projekt *Bicycle Observatory* wurden die Daten, die für das Projekt relevant sind, sukzessive in den DMP integriert. Der DMP bildet immer den aktuellen Ist-Bestand der im Projekt integrierten und aggregierten Daten ab. Daher muss die Pflege des DMP kontinuierlich durchgeführt werden.

Für die Pflege des DMP hat sich im Projektverlauf herausgestellt, dass ein Partner bzw. Projektmitarbeiter/-in die Hauptverantwortung für den DMP trägt. Dieser erfasst bei neuen Datensätzen die ihm grundlegend bekannten Metadaten. Für zusätzliche Informationen werden dann weitere Projektmitarbeiter/-innen herangezogen und die notwendigen Informationen nachgefragt. Die *Confluence*-basierte Umsetzung des DMP bringt den Vorteil, dass jegliche Änderungen und Ergänzungen auf der Ebene eines Datensatzes einfach dokumentiert und nachverfolgt werden (Versionierung der einzelnen Seiten). Werden Zusatzinformationen zu den Datensätzen gegeben, die nicht direkt in das Format der Datensatzbeschreibung passen oder erst mit dem Verantwortlichen für das Datenmanagement diskutiert werden müssen, kann die Kommentarfunktion bei den einzelnen Seiten von *Confluence* genutzt werden.

## 4 Zusammenfassung und Ausblick

Im Projekt *Bicycle Observatory* wurde das Thema Datenmanagement und die Umsetzung des dafür notwendigen DMP von Projektstart an mitgedacht. Als zentraler Punkt wurde neben der Erfüllung der notwendigen Kriterien der DSGVO auch ein einheitliches und möglichst standardisiertes Vokabular für das Metadaten-Management identifiziert. Das *OGD Metadaten 2.4* Schema bietet hier eine solide Grundlage, dass in den optionalen Feldern auch nützliche Definitionen für die Abbildung von räumlichen Datensätzen bietet.

Die Erstellung und Pflege eines DMP muss als ein integraler Bestandteil eines Forschungsprojekts gesehen werden. Einerseits benötigt die Pflege einen in den Projektablauf eingebetteten Prozessablauf, auf der anderen Seite ist auch die notwendige Software Voraussetzung dafür. Damit das Datenmanagement im Projekt auch gelebt wird, muss mit der Software eine einfache und reibungslose Pflege des DMP möglich sein. In *Bicycle Observatory* wurde dazu die Plattform *Confluence* gewählt, da diese bereits aus Vorgängerprojekten für den Zweck der Projektdokumentation bekannt ist. Für die einfachere Pflege des DMP wurde dazu eine eigene DSGVO- und *OGD Metadaten 2.4*-konforme Confluence-Vorlage erstellt. Diese Vorlage wurde im Rahmen der Bearbeitung im Projekt laufend optimiert und wird als Projektergebnis auf Github<sup>6</sup> unter der CC BY 4.0 als OpenData zur Verfügung gestellt.

In zukünftigen Projektphasen/Projekten wäre es erstrebenswert, durch ein zunehmendes Bewusstsein für das Thema des Datenmanagements und die einfachere Umsetzung des DMPs eine dezentralere Pflege des DMPs zu erreichen. Datenmanagement sollte direkt in den Organisationen stattfinden wo das größte Wissen über die Datensätze vorliegt.

## Danksagung

Das Projekt Bicycle Observatory – Am Puls des Radverkehrs (FFG Nr. 865176) wird im Rahmen des Programms „Mobilität der Zukunft“ vom österreichischen Bundesministeriums für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie (BMK) gefördert.

Am Projekt wirken die folgenden Konsortialpartner mit: (1) Universität Salzburg, Interfakultärer Fachbereich für Geoinformatik – Z\_GIS (Koordination), (2) Salzburg Research Forschungsgesellschaft mbH, (3) PRISMA solutions EDV-Dienstleistungen GmbH, (4) Helios.

## Literatur

- Brocza, U., & Kollarits, S. (2020). Dashboard Radverkehr: Alles im Blick. Das Lagebild zum Radverkehr – vom Überblick zu den Details und zurück. *AGIT – Journal für Angewandte Geoinformatik*, 6-2020, 238–243.
- Cooperation OGD Österreich (2017). *OGD Metadaten Österreich 2.4*. White Paper. Retrieved January 31, 2020, from [https://www.ref.gv.at/fileadmin/user\\_upload/OGD-Metadaten\\_2-4\\_20170228.pdf](https://www.ref.gv.at/fileadmin/user_upload/OGD-Metadaten_2-4_20170228.pdf).
- DAMA International (2017). *DAMA-DMBOK*. (D. Henderson & S. Earley, Eds.) (2nd Ed.). Technics Publications.
- DCAT-AP (2020). *About DCAT Application Profile for data portals in Europe*. Retrieved February 03, 2020, from <https://joinup.ec.europa.eu/solution/dcat-application-profile-data-portals-europe/about>.
- digitales.oesterreich (2020). *Public-Sector-Information-Richtlinie*. Retrieved February 03, 2020, from <https://www.digitales.oesterreich.gv.at/public-sector-information-richtlinie>.
- Dove, E. S. (2018). The EU General Data Protection Regulation: Implications for International Scientific Research in the Digital Era. *The Journal of Law, Medicine & Ethics*, 46(4), 1013–1030.

---

<sup>6</sup> [https://github.com/BiObs/Confluence-Vorlage\\_Datenmanagementplan](https://github.com/BiObs/Confluence-Vorlage_Datenmanagementplan)

- EC (2020). *Horizon 2020 Online Manual Data Management*. Retrieved February 03, 2020, from [https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management\\_en.htm](https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm).
- EU (2016). *Richtlinie (EU) 2016/680 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten durch die zuständigen Behörden zum Zwecke der Verhütung, Ermittlung, Aufdeckung oder Verfolg*. Retrieved February 03, 2020, from [https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=uriserv%3AOJ.L\\_.2016.119.01.0089.01.DEU](https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=uriserv%3AOJ.L_.2016.119.01.0089.01.DEU).
- ICPSR (2012). *Guidelines for Effective Data Management Plans*. Retrieved February 03, 2020, from <https://www.icpsr.umich.edu/files/datamanagement/DataManagementPlans-All.pdf>.
- Inspire (2020). *Inspire Österreich*. Retrieved February 03, 2020, from <https://www.inspire.gv.at/>.
- Loidl, M., Wagner, A., Kaziyeva, D., & Zagel, B. (2020). Bicycle Observatory – eine räumlich differenzierte, kontinuierliche Beobachtung der Fahrradmobilität. *AGIT – Journal für Angewandte Geoinformatik*, 6-2020, 263–271.
- Miller, H. J. (2017). Geographic information science I: Geographic information observatories and opportunistic GIScience. *Progress in Human Geography*, 41(4), 489–500.
- VerbundFDB (2015). *Checkliste zur Erstellung eines Datenmanagementplans in der empirischen Bildungsforschung*. Retrieved February 03, 2020, from [https://www.forschungsdaten-bildung.de/files/fdbinfo\\_2.pdf](https://www.forschungsdaten-bildung.de/files/fdbinfo_2.pdf).