

# Empfehlungen für die Datenbereitstellung im mFUND (Merkblatt für Fördernehmer)



Bundesministerium  
für Digitales  
und Verkehr

**mobilithek**

**Stand: Version 1.6 –27. Juli 2022**

## **Schritt 1: Auswählen**

### ***Welche Daten kommen für eine Bereitstellung als Open Data in Frage?***

Unter Open Data versteht man grundsätzlich solche Daten, deren Inhalte von allen frei und zu jedem Zweck verwendet, verändert und geteilt werden können<sup>1</sup>.

Bei der Datenbereitstellung im mFUND geht es grundsätzlich nur um solche Daten, die im Rahmen des aus dem mFUND geförderten Projektes entstanden sind. Selbstverständlich steht es jedem frei, weitere Daten mit Bezug zum Projekt oder davon unabhängige Daten zusätzlich zur Verfügung zu stellen.

Die bereitzustellenden Daten sollen die grundsätzlichen Open-Data-Eigenschaften<sup>2</sup> erfüllen:

- Elektronisch gespeichert  
*Die Daten liegen bereits elektronisch gespeichert vor. Zeichnungen, Notizen, jegliches „analoges“ Material ist nicht gemeint. Eine Digitalisierung ausschließlich zum Zwecke der Datenbereitstellung ist nicht erforderlich.*
- Maschinenlesbar  
*Es sollen nur solche Daten bereitgestellt werden, die maschinenlesbar, also automatisch interpretierbar und analysierbar sind. Nicht gemeint sind in diesem Sinne also Dokumente (reine Schriftstücke, E-Mails etc.), Fotos oder Dateien, die in nicht weiter bearbeitbaren Formaten vorliegen (z.B. PDF-Dokumente). Letztere können jedoch als begleitendes Material zusätzlich zu den maschinenlesbaren Daten bereitgestellt werden.*
- In Sammlungen strukturiert  
*Wenig hilfreich für Dritte sind völlig unstrukturierte Einzeldaten (z.B. einzelne Messwerte, deren Zusammenhang nicht erkennbar ist). Daher sind nur solche Daten weiter zu betrachten werden, die bereits in einem Sinnzusammenhang strukturiert sind.*
- Standardisierte und offene Formate  
*Neben der Maschinenlesbarkeit ist die Standardisierung eines Formats wichtig. Die Daten sollten in einem Format vorliegen, das in Form eines offen und unentgeltlich nutzbaren Standards präzise definiert und dokumentiert ist. Offene Standards garantieren, dass Daten in diesem Format jederzeit und von jedem korrekt verarbeitet werden können. Zudem sollte statt eines proprietären Formats nach Möglichkeit immer ein offenes gewählt werden, damit die Weiterverarbeitung der Daten nicht an das Vorhandensein einer möglicherweise speziellen und teuren Software gebunden ist. Beispiele für solche Formate sind: CSV, XLSX, ODS, XML, JSON, RDF, TXT, GeoJSON, KML etc.*

Alle Daten, die die vorgenannten Bedingungen (1. Auswählen) erfüllen, sollten im weiteren Verlauf einer Prüfung (2. Prüfen) unterzogen werden. Folgende Daten sind zu betrachten:

---

<sup>1</sup> Siehe „The Open Definition“: <https://opendefinition.org/>

<sup>2</sup> Einen guten Überblick über die Prinzipien von offenen Daten gibt die Konrad-Adenauer-Stiftung unter [https://www.kas.de/documents/252038/253252/7\\_dokument\\_dok\\_pdf\\_44530\\_1.pdf](https://www.kas.de/documents/252038/253252/7_dokument_dok_pdf_44530_1.pdf). Der Fokus liegt in dieser Beschreibung auf „Verwaltungsdaten“, aber die Grundsätze lassen sich weitestgehend auch auf private Daten übertragen.

- **Unbearbeitete Daten**  
*Für den Zweck der Bereitstellung als Open Data sind unbearbeitete Daten („Rohdaten“) besonders interessant, da diese noch keine Bearbeitung und Interpretation enthalten. Ein Beispiel hierfür sind Sensor-Messdaten. Grundsätzlich sind solche Daten zu veröffentlichen. Manchmal handelt es sich dabei jedoch um sehr große Datenmengen, die zudem stark fehlerbehaftet sein können. Daher ist alternativ die Veröffentlichung von aufbereiteten, plausibilisierten Daten zu prüfen.*
- **Aufbereitete, plausibilisierte Daten**  
*In vielen Fällen werden Rohdaten erst einer Qualitätsprüfung, Fehlerkorrektur und Strukturierung unterzogen, bevor sie weiterverarbeitet werden. Für viele Nutzer sind solche Daten interessanter, als die unbearbeiteten Daten, da sie ohne weitere Vorverarbeitungsschritte nutzbar sind.*
- **Ergebnis- und Analysedaten**  
*Bei der Verarbeitung von Daten entstehen oftmals verschiedene Zwischenstände, die beispielsweise nach einer Kalibrierung oder Validierung anfallen. Auch durch die Veredelung von Daten, bspw. durch Kombination mit anderen Daten, durch statistische Auswertung oder Interpretation können neue Datenprodukte entstehen. Auch für diese ist eine Veröffentlichung zu prüfen.*

## **Schritt 2: Prüfen**

### ***Worauf sollte vor einer Veröffentlichung von Daten unbedingt geachtet werden?***

Einige Eigenschaften von Daten können einer Veröffentlichung im Sinne von Open Data entgegenstehen. Um insbesondere den Datenschutz, aber auch andere rechtliche Aspekte zu wahren, sind die Daten vor einer möglichen Veröffentlichung mindestens auf folgende Aspekte hin zu prüfen (ggfs. wird diese Liste erweitert):

- **Beinhalten die Daten personenbezogene oder personenbeziehbare Daten im Sinne der Datenschutzgrundverordnung (DSGVO) und des Bundesdatenschutzgesetzes (BDSG)?**  
*Hierzu zählen alle Informationen, die einen Rückschluss auf natürliche Personen zulassen. Manchmal ist dieser Rückschluss erst durch die Kombination mit anderen Daten möglich. Daher ist zu prüfen, ob bei den zu veröffentlichenden Daten diese Gefahr besteht. Hier kann nur nach bestem Wissen und Gewissen entschieden werden, da niemand einen vollständigen Überblick über mögliche Kombinationsmöglichkeiten mit anderen Daten besitzt. Sollte aber ein Personenbezug in den Daten vorliegen oder erahnt werden, so ist zu prüfen, ob durch geeignete Maßnahmen der Personenbezug entfernt werden kann. Hierzu zählt die Anonymisierung, beispielsweise durch Aggregation der Daten oder durch Löschung einzelner Datenelemente. Grundsätzlich kann auch die Einwilligung von Betroffenen für eine Veröffentlichung eingeholt werden. In diesem Fall muss sichergestellt sein, dass eine solche Einwilligung von den Betroffenen jederzeit widerrufen werden kann.*
- **Würde eine Veröffentlichung der Daten das geistige Eigentum, Urheberrechte Dritter oder vertragliche Bedingungen verletzen?**  
*Wenn durch eine Veröffentlichung von Daten Urheberrechte Dritter, Betriebs- und Geschäftsgeheimnisse, Marken- oder Patentrechte verletzt würden, ist zu prüfen, ob eine*

*Einwilligung der Betroffenen eingeholt werden kann. Andernfalls können solche Daten nicht veröffentlicht werden. Gleiches gilt, wenn Daten aufgrund von eingegangenen Verträgen nicht weitergegeben werden dürfen.*

- Handelt es sich um vertrauliche Daten?  
*Gibt es bei den Daten eine Geheimhaltungs- oder Vertraulichkeitspflicht oder wurden die Daten vertraulich erhoben oder übermittelt und besteht das Interesse Dritter an einer vertraulichen Behandlung fort, gilt das gleiche wie bei personenbezogenen Daten: Anonymisieren, Einwilligung einholen oder Veröffentlichung unterlassen.*
- Würde eine Veröffentlichung der Daten den Schutz öffentlicher Belange verletzen?  
*Es kann vorkommen, dass sich die Veröffentlichung von Daten negativ auf sog. öffentliche Belange auswirkt. Hierzu zählt beispielsweise die öffentliche Sicherheit, die dadurch gefährdet würde, wenn detaillierte Informationen veröffentlicht würden. Ein Beispiel hierfür sind detaillierte Daten über kritische Infrastrukturen. Ein weiteres Beispiel wäre, wenn durch eine Datenveröffentlichung laufende Gerichtsverfahren beeinträchtigt würden. Sollte ein solcher Fall bei den zu veröffentlichenden Daten zu erwarten sein, ist entweder eine Reduktion des Informationsgehalts zu prüfen oder eine Veröffentlichung zu unterlassen.*
- Hat eine Veröffentlichung der Daten negative Auswirkungen auf das eigene Geschäftsmodell?  
*In mFUND-Projekten geht es um die Entwicklung datenbasierter Innovationen. Dabei können Daten entstehen, die fester Bestandteil einer im Rahmen des Projektes entwickelten Geschäftsidee sind. Wenn eine Veröffentlichung negative Auswirkungen auf die Tragfähigkeit dieser Idee hätte, sollte eine Reduktion des Informationsgehalts vor Veröffentlichung in Betracht gezogen oder ganz auf eine Veröffentlichung verzichtet werden.*

Sollten Belange gegen eine direkte Veröffentlichung sprechen, sind die Daten wenn möglich soweit zu bearbeiten, damit eine Veröffentlichung möglich wird. Der Umfang nicht veröffentlichter Daten ist auf ein Minimum zu reduzieren.

**Final nicht zur Veröffentlichung geeignete Daten sind unter Angabe von Gründen nachvollziehbar mit den Zwischen- und Abschlussberichten des Projektes zu erläutern.**

### **Schritt 3: Veröffentlichung vorbereiten**

#### ***Wo können die Inhaltsdaten gespeichert werden?***

Für die Veröffentlichung von Open Data sind sowohl die Daten selbst (**Inhaltsdaten**), als auch eine Beschreibung der Daten (**Metadaten**) bereitzustellen. Daher sollte nach der Prüfung (Schritt 2) die Veröffentlichung vorbereitet werden. Ein Ziel des mFUND ist es, möglichst viele Daten aus den Projekten zur Weiterverwendung und weiteren Wertschöpfung verfügbar zu machen. Als zentraler Zugangspunkt zu (offenen) Daten aus allen Themenbereichen der Mobilität dient die Mobilithek des BMDV: [www.mobilithek.info](http://www.mobilithek.info). Die Mobilithek ist durch die mFUND-Projekte primär für die Veröffentlichung ihrer Daten zu nutzen. Hierfür muss zunächst für die zu veröffentlichenden Daten eine zentrale Frage beantwortet werden können: Gibt es im Projekt (bzw. im Unternehmen oder in der Institution) Möglichkeiten, die **Inhaltsdaten** an öffentlich zugänglicher Stelle längerfristig abzulegen?

- Ja, weil z.B. ein eigener Webserver zur Verfügung steht oder eine öffentliche Cloud-Infrastruktur genutzt werden kann.  
*Es kann direkt mit Schritt 4 (Metadaten erfassen) fortgefahren werden.*
- Nein, eine entsprechende Infrastruktur steht nicht zur Verfügung.  
*Die Mobilithek dient dazu, statische und dynamische Daten zur Mobilität zu verteilen. Insofern verfügt die Mobilithek einerseits über standardisierte Schnittstellen, über die dynamische Daten (bspw. Echtzeitdaten) Dritten im Sinne eines Abonnements über einen sog. „Broker“ zur Verfügung gestellt werden können. Andererseits können auch Dateien in den Speicher der Mobilithek hochgeladen und dort zum Download angeboten werden. Dieser Upload ist vor allem für offene statische Daten (d.h. Daten, die sich nach der Erhebung nicht mehr verändern) die geeignetste Variante. Es ist dabei jedoch zu beachten, dass die Mobilithek als kostenfreies Angebot des BMDV eine Limitierung in Bezug auf die Dateigröße enthält: Der Upload in den Speicher ist auf 10 GB pro Datei begrenzt. Das Limit für dynamische Daten liegt bei 100 MB pro Datei.*

### **Schritt 4: Daten veröffentlichen**

#### ***Worauf ist bei der Veröffentlichung von Meta- und/oder Inhaltsdaten zu achten?***

Ihre Daten veröffentlichen Sie auf der Mobilithek in Form von sogenannten Datenangeboten oder Publikationen. Alle wichtigen Informationen und Antworten auf Fragen zur Erstellung von Datenangeboten erhalten Sie in den FAQs der Mobilithek: <https://mobilithek.info/help/FAQ>. Sollte Ihre Frage dort nicht beantwortet werden, können Sie das Kontaktformular (<https://mobilithek.info/contact>) benutzen oder sich auch an den Servicedesk (Hotline) wenden: [OS-SD-Mobilithek@o-s.de](mailto:OS-SD-Mobilithek@o-s.de) oder Hotline-Nr.: +49 69 689702810. Der Servicedesk (Hotline) ist zu folgenden Zeiten erreichbar: Montag - Freitag: 07:00 - 19:00 Uhr, Samstag: 08:00 - 14:00 Uhr.

Sollte Ihre Organisation über ein eigenes Metadatenportal verfügen, ist es grundsätzlich auch möglich, dieses mit der Mobilithek so zu verknüpfen, dass die Datenangebote aus Ihrem Metadatenportal in der Mobilithek dargestellt werden. Eine nochmalige Erfassung von Metadaten in der Mobilithek selbst erübrigt sich dann. Wenn Sie hierzu Fragen haben, wenden Sie sich bitte ebenfalls an den Servicedesk oder nutzen Sie das Kontaktformular.

### **Schritt 5: Berichten**

#### ***Was soll dokumentiert werden?***

Bitte berichten Sie in Ihren **Zwischen- und Abschlussberichten** über den Fortschritt der Datenbereitstellung in Ihrem Projekt. Hierbei sollte insbesondere auch beschrieben werden, welche Daten nicht zur Veröffentlichung geeignet sind und worin die Gründe hierfür liegen.