

# **GUIDO – Open Data Process Guide für Ministerien**

## **Empfehlung zur Umsetzung, Weiterentwicklung und Einbettung in den Kontext weiterer Maßnahmen**

---

### *Einleitung & Kontext*

Im 12-wöchigen Fellowship hat ein vierköpfiges Team aus dem Tech4Germany-Umfeld, mit Unterstützung von GovData und dem Auswärtigen Amt mit dem Open Data Process Guide ein erstes MVP (Minimum Viable Product) entwickelt, welches Ministerien dabei unterstützt den Prozess von der Datenidentifizierung bis zur Bereitstellung auf GovData zu vereinfachen. Da die Lösung innerhalb kurzer Zeit entwickelt wurde und das Team dieses nach Ende des Fellowship-Programms vollständig übergibt, entstehen mehrere Herausforderungen.

Da *GUIDO* zwar am Beispiel der aus Interviews gewonnenen Erkenntnisse aus dem Auswärtigen Amt entwickelt wurde, aber aufgrund der kurzen Projektdauer dort nicht implementiert werden konnte, gibt es noch keine realen Erfahrungswerte bezüglich der Implementierung innerhalb eines Ministeriums. Es ist daher davon auszugehen, dass vor, während und nach der Implementierung des Tools Fragen seitens der jeweiligen Ministerien entstehen. Der Umsetzungsplan enthält daher einen **Fahrplan für Ministerien**, der grundlegende Fragen beantwortet, aufzeigt, wie eine erfolgreiche Implementierung umzusetzen ist und wer die Ansprechpartner sind.

Des Weiteren enthält der Umsetzungsplan **strategische Überlegungen**, insbesondere im Bezug auf die Integration GUIDOs in bereits laufende Planungen des BMI zum IT-Unterstützungstool.

Zuletzt gibt es Arbeitsschritte, die notwendig sind, um GUIDO **vom MVP zu einem vollumfänglichen Tool** zu entwickeln. Darüber hinaus enthält dieses Dokument eine Vielzahl an Überlegungen, wie das Tool zukünftig weiterentwickelt werden kann, um es noch nutzerfreundlicher zu machen, den Prozess weiter zu vereinfachen und die Datenqualität automatisiert zu erhöhen. Diese **Empfehlungen zur Weiterentwicklung von GUIDO** sind ebenfalls Teil des Umsetzungsplans.

## Vom MVP zum vollumfänglichen Tool

Da der erste Teil des Fellowships darin bestand, die Problemlandschaft durch Nutzer- und Experten-interviews aufzuzeigen und sich darin für eine möglichst sinnvolle Lösung zu entscheiden, konnte das Programmieren einer konkreten Lösung erst nach Abschluss dieser Phase beginnen. Insofern gaben wir uns die Entwicklung eines MVPs als eine realistische Zielsetzung (im Kontrast zu einem vollumfänglichen Tool).

Im folgenden wird die Grundarchitektur beschrieben um dann auf die Punkte einzugehen, die noch ausgebaut werden müssten auf dem Weg von einem MVP zu einem fertigen Produkt.

### Architekturbeschreibung

Eine der Grundideen ist, eine möglichst flexible Definition der Prozesse zur Bereitstellung von Open Data zuzulassen. Der Grund hierfür ist, dass anzunehmen ist, dass sich die Gegebenheiten in den Ministerien unterscheiden, was Ausgestaltungen des Prozesses angeht. Insbesondere Verantwortlichkeiten wie die rechtliche Prüfung oder die Veröffentlichung auf der Ministeriumswebseite werden unterschiedlich organisiert sein. Auch innerhalb eines Ministeriums wollten wir die Möglichkeiten offenlassen, dass bspw. ein Referat eine andere "Vorschleife" drehen muss oder kann, bevor sie sich auch in den als Standard definierten Prozess des Ministeriums einklinkt. Aus dieser Flexibilitätsanforderung heraus haben wir uns für eine komplette Beschreibung des Prozesses im text-basierten JSON-Format entschieden. Ein Prozess besteht dabei aus einer chronologisch sortierten Abfolge von Modulen und optionalen Verantwortlichkeiten pro Modul. Jedes Modul hat Name, Beschreibung und dann eine Reihe an Feldern die ausgefüllt werden können oder müssen. Auch diese Modul-Beschreibung findet im gleichen JSON-Dokument statt. Nach einer minimalen Einarbeitung ist es also möglich, Prozesse (bestehend aus Modulen) komplett textuel zu definieren, ohne Programmierkenntnisse. Sobald ein Prozess zur Datenbereitstellung gestartet wird, nennen wir es eine "Bereitstellung" und die darin instanziierten Module "Aufgaben" mit "Feldern". Wie weiter unten beschrieben, kann eine Prozessdefinition auch von einem BPMN Diagramm importiert werden.

Weiterführende technische Details können dem Readme im GitHub Repository entnommen werden: <https://github.com/tech4germany/open-data-process-guide>.

## Was noch zu implementieren ist

- Ein paar Funktionen wurden zum Ende hin recht pragmatisch umgesetzt um sie in der Demo zeigen zu können. So werden an manchen Stellen im Code bestimmte Felder von Modulen fest angenommen. Dies läuft allerdings dem Prinzip der kompletten Definition ausgehend von einer JSON-Datei entgegen. Hier sollte ein Refactoring angegangen werden, dass diese Logik wieder intakt macht.
- Es sind nicht alle Modul-Felder, die in der JSON-Datei vorkommen können ausimplementiert. Hier ist es nötig für ein paar weitere Feld-Typen die Darstellung als auch das dahinter liegende Datenmodell in die restliche Architektur zu integrieren.
- Obligatorische Felder sind zwar als solche definiert und werden entsprechend dargestellt, aber die Logik dies sicherzustellen ist noch nicht vorhanden.
- Zusätzliche Info-Felder die im Click-Dummy rechts neben den Formularfeldern erscheinen, bspw. im Schritt der rechtlichen Prüfung, werden noch nicht dargestellt. Diese sollten auch in der config.json Datei abgelegt werden und entsprechend gerendert werden.
- Die Persistierung von Prozessdefinitionen und Bereitstellungen erfolgt via Lists in SharePoint. Diese sind ähnlich wie Tabellen aufgebaut. Um den Aufwand einer Anpassung an diese Datenstruktur gering zu halten, haben wir sehr pragmatisch einfach die jeweiligen Daten pro Prozessdefinition oder pro Bereitstellung als serialisiertes JSON in ein Datenfeld in der entsprechenden List geschrieben. Ein sauberes Vorgehen wäre hier sicherlich die Datenstruktur mehr so zu nutzen wie sie vorgesehen ist - oder alternativ eine andere Datenbanklösung anzuschließen.
- Der Import von BPMN wurde nur prototypisch umgesetzt. Um alle BPMN Symbole zu unterstützen (bspw. parallele Abläufe) wäre noch eine nicht unerhebliche Anstrengung beim Parsen der *.bpmn* Datei notwendig. Parallel Abläufe sind auch in der Architektur bisher nicht vorgesehen. Auch wäre für komfortables Arbeiten mit BPMN als Startpunkt der Prozessdefinition ein flexibleres Mapping von BPMN-Task auf Module-Name angenehm. Bspw. eine Art Import-Assistenten der anhand "Meinen Sie das damit"-Fragen Ambiguitäten auflösen kann. Auch beim Import der Verantwortlichkeiten (sog. "Lanes" in BPMN) wäre ein komfortables Mapping von Titel zu tatsächlichen Personen mit Email-Adressen ein wichtiger Schritt.
- Die Bedingungen unter welchen die im Bereitstellungsprozess befindlichen Daten geändert werden können, bspw. sollten, bevor sie der Veröffentlichung übergeben werden, müsste noch besser definiert werden. Aktuell kann jeder mit

Schreibrechten im entsprechenden Dokumente-Ordner auf SharePoint bearbeiten. Dies müsste ggf. feingranularer geregelt werden, auch mit einem Editing-Lock-Feature nachdem bspw. eine rechtliche Prüfung stattgefunden hat.

- Auf die Nutzerverwaltung wurde im MVP kein Fokus gelegt. Hier müsste an mehreren Stellen nachgefasst werden um bspw. sicherzustellen das auch wirklich nur die Verantwortlichen die zur Bearbeitung einer Aufgabe eingeladen werden, diese bearbeiten dürfen. Aktuell könnte jeder mit der entsprechenden URL Parameter Encodierung diese Aufgaben übernehmen.
- Aktuell können Aufgaben zu denen Verantwortliche spezifiziert werden auch vom aktuellen Bearbeiter übernommen werden anstatt die entsprechende Stelle zu benachrichtigen. Diese Flexibilität müsste ggf. nochmal überdacht werden und einschränkbar sein. So ist durchaus vorstellbar, dass es Aufgaben gibt, die zwingend delegiert werden müssen.
- Der Lebenszyklus der Dateien im SharePoint-Dokumente-Ordner muss durchdacht werden. Aktuell wird keine Löschung am Ende des Datenbereitstellungsvorgangs vorgenommen. Dies führt aber zu einer doppelten Datenhaltung der Dateien auf SharePoint und der veröffentlichten Dateien auf der Ministeriumsseite. Hier müsste ein schlüssiges Konzept erarbeitet werden.
- Der letzte Arbeitsschritt in jeder Bereitstellung, nämlich die Übergabe an GovData, wurde nicht angegangen. Die Aufgabe besteht hier darin, die als JSON vorliegenden Schlüssel-Wert-Paare, die sich aus den ausgefüllten Feldern ergeben, als XML/RDF-Datei auszugeben, die dem DCAT-AP.de Standard folgt. Hierbei würden alle abgeschlossenen Bereitstellungen in eine Katalog-XML Datei zusammengeführt werden. Diese Datei muss dann öffentlich zugänglich sein, so dass GovData sie harvesten, das heißt automatisiert im System einlesen, kann. Hier ist es ratsam, das Vorgehen eng mit GovData abzustimmen um von deren reichhaltiger Erfahrung in diesem Arbeitsschritt zu profitieren.

### *Was bedeutet das für mich als Ministerium*

GUIDO ist nach den oben beschriebenen Punkten also nicht als sofort einsatzfähig zu betrachten. Es ist noch Arbeit nötig um das Tool prinzipiell einsetzen zu können, noch bevor man sich an die Ministeriums-spezifische Anbindung und Integrierung macht. Da der Code öffentlich auf GitHub ist, würde hier allerdings ein einmaliger Aufwand ausreichen und allen anderen zu Gute kommen die GUIDO einsetzen wollen. Wir empfehlen, unsere Design-Artefakte (im Click-Dummy) bei einem Ausbau der

Implementierung im Blick zu haben, da die bisherige Implementierung ein Ausfluss derselbigen war und sie die Bündelung der Nutzerinterviews und mehrere Iteration hinsichtlich der Nutzbarkeit und Nutzerfreundlichkeit eingeflossen sind.

## *Weiterentwicklung von GUIDO*

Im Folgenden werden mehrere Features erläutert, welche aufgrund von Zeitgründen nicht im Rahmen des Projekts umgesetzt werden konnten. Die Features sind dabei zur besseren Übersicht und Einschätzung in drei grobe Kategorien unterteilt - **Big Hitter, Quick Wins und Low Hanging Fruits**. Sie variieren in den Dimensionen geschätzter Aufwand und erwarteter Nutzen.

Die Ableitung der Features und die Einschätzungen hinsichtlich Aufwand und Nutzen basieren dabei auf den Erkenntnissen des Projektteams aus der Research-Phase (Identifikation von Best Practices, Analyse anderer Open Data Portale, etc.) und den Nutzerinterviews, die zu Beginn explorativ geführt wurden und später gezielt genutzt wurden, um den MVP zu vertesten.

### *Big Hitter (Hoher Aufwand, hoher Nutzen)*

#### **Data Refinement Modul:**

Eine wertvolle Ergänzung würde die Aufwertung der Datenqualität darstellen. Je nachdem, in welchem Zustand die Rohdaten vorliegen, kann es beispielsweise Sinn machen im Fall von Tabellen bestimmte Spalten zusammenzuführen, zu trennen, oder Inhalte von unnötigen Satzzeichen zu bereinigen. Abhängig davon, aus welchen Systemen die Daten stammen, können auch Kürzel vorliegen, die nur im entsprechenden Kontext bekannt sind. Ein Beispiel dafür wären Abkürzungen für Regionen, die außerhalb des Kontextes nicht bekannt sind. Hier bietet sich an, die entsprechenden Werte auszusprechen, oder DCAT-AP.de-konforme regionale Identifier zu verwenden. Prinzipiell wäre es wünschenswert alle Schritte einzuleiten, die die Daten für einen Linked Open Data aufbereitet und verwendbar macht. Je nachdem, in welchen Pipelines und Systeme die Daten landen, ist dort anschließend weniger Aufwand nötig, um sie tatsächlich nutzbar zu machen. Bei der Planung und Umsetzung eines solchen Moduls würde sich anbieten entsprechende Experten aus der Wissenschaft und der Open Data Community zu Rate zu ziehen. Die ist vor allem deshalb empfehlenswert,

weil es in diesem Feld viele Standards und Best Practices gibt und aktive Forschung dazu stattfindet. Hinsichtlich der Verwendung oder Einbindung bestehender Tools könnte sich beispielsweise OpenRefine (<https://openrefine.org/>) und FME (<https://www.safe.com/fme/>, für Geodaten) anbieten, diese wurden uns bereits in Experteninterviews genannt. Da es sich um ein dynamisches Feld handelt mit potentiell vielen spezialisierten Funktionen (bestimmte Datenveredelungsschritte machen ggf. nur für sehr spezifische Bereiche Sinn), würde es sich hier auch anbieten, eine Vorschlagsfunktion einzubauen, mit Hilfe derer Datenbereitsteller Funktionalitäten aus ihrem Fachbereich heraus vorschlagen könnten. Beispielhaft für eine solche Funktion wäre, dass jedesmal, wenn Daten vom Typ X bereitgestellt werden, vorher eine Y-Prüfung darauf angewendet wird. Idealerweise kann das Modul möglichst viele Vorschläge automatisch treffen und der Datenbereitsteller muss die Vorschläge letztendlich nur noch akzeptieren oder ablehnen. Ebenfalls wäre es auch denkbar, dass das Modul eine Rückfrage an den User stellt, um eine gewissen Vorauswahl zu treffen und anschließend eine Handlungsempfehlung zu geben ("Handelt es sich bei dieser Spalte um Geodaten? Dann schlage ich Folgendes vor..."). Eine weitere Überlegung könnte sein, aus dem Modul heraus eigene oder externe Experten anzufragen, um Einblick darin zu erhalten, wie sie den jeweiligen Datensatz aufbereiten würden.

### **Automatisierte Datenanonymisierung:**

Enthalten Dateien personenbezogene Daten, können diese vorerst nicht für eine Veröffentlichung herangezogen werden. Allerdings können diese Dateien dennoch wertvolle Informationen enthalten und nützlich für Forschung und Wirtschaft oder interessant für Bürger und Bürgerinnen sein, sobald die personenbezogenen Details entfernt wurden. Daher ist es wichtig, personenbezogene Daten zu identifizieren und gegebenenfalls entfernen zu können. Diesen Anspruch könnte ein Feature innerhalb von GUIDO abdecken. Eine sinnige Platzierung wäre aus unserer Sicht im Prozessschritt "Prüfung auf rechtliche Einschränkungen". In diesem Schritt werden die Dateien auf die Gesetzesgrundlagen geprüft. Ein Teil dieser Prüfung beinhaltet die Kontrolle, ob die Dateien personenbezogene Daten enthalten, wie etwa Namen, Adressen oder Telefonnummer. An dieser Stelle könnte es einen Button geben, durch welchen eine automatische Identifizierung personenbezogener Daten stattfindet. Anschließend wird für den:die Bearbeiter:in der betreffenden Bereich in der Datei hervorgehoben und es kann bewertet werden, ob eine Löschung der entsprechenden Details sinnvoll ist, oder ob die ganze Datei aus dem Bereitstellungsprozess entfernt werden muss. Auch ist es

möglich dieses Feature als generelles Überprüfungsfeature anzubieten, welches ohne einwirken des:der Nutzer:in schon beim Upload der Dateien im Hintergrund läuft und meldet, sobald personenbezogene Daten gefunden wurden. Dies würde dem:der Bearbeiter:in Gewissheit geben, dass die Dateien keine sensiblen Daten enthalten, solange keine entsprechende Meldung angezeigt wird. Egal in welche Ausprägung, der Schritt der manuellen Kontrolle der Daten würde entfallen und der Prozess der Überprüfung wird dem:der Ministeriumsmitarbeiter:in erleichtert. Ebenfalls kann so auch langfristig auch die Angst, sich an der Open Data Bereitstellung zu beteiligen, vermindert werden. Das Anonymisierungstool [openredact](#) des Prototyp Funds, gefördert durch das Bundesministerium für Bildung und Forschung, könnte hier als Vorbild dienen. Auch könnte man eine direkte Zusammenarbeit anstreben.

### *Quick Wins (Niedriger Aufwand, moderater Nutzen)*

#### **User Interface für die Prozessdefinition:**

Damit jedes Ministerium den Prozess individuell an seine vorhandenen Ressourcen (zuständige Mitarbeiter:innen und Abteilungen) anpassen kann, wird innerhalb von GUIDO ein nutzerfreundliches Interface benötigt, welches die notwendigen Funktionen dafür zu Verfügung stellt. Zuerst soll es möglich sein eine BPMN-Datei, welche man zuvor in einem Prozessmanagement Tool (z.B. Signavio) erstellt hat, über ein Upload-Feld in GUIDO hochzuladen. Somit sollten optimalerweise die wichtigsten Prozessschritte bereits definiert sein und in einer Vorschau angezeigt werden.

Dennoch soll es darüber hinaus möglich sein, über GUIDO Anpassungen vorzunehmen, dazu gehören:

- Löschen eines Prozessschrittes
- Hinzufügen eines Prozessschrittes
- Ändern der Reihenfolge der Prozessschritte
- Zuständige Bearbeiter:in eines Prozessschrittes ändern

Anschließend sollte dieser Prozess als Standard für jede Bereitstellung im jeweiligen Ministerium gelten. Die Berechtigung für die Funktion den Prozess zu definieren sollte nur einer berechtigten Person möglich sein (z. B. der:die Open Data Beauftragte:r im Ministerium).

#### **Nachträgliche Bearbeitung von Datensätzen:**

Eine weitere wichtige Funktion wäre, abgeschlossene Bereitstellungsvorgänge, die bereits auf GovData veröffentlicht wurden, im Nachhinein bearbeiten zu können. Da der RDF-Output nicht mehr implementiert wurde, wurde auch dieses Feature ausgeklammert. Hier gilt es, die entsprechenden Identifier in der RDF Datei konsistent zu halten, um Datensätze entsprechend als Updates zu flaggen, damit das GovData Harvesting-System diese als solches erkennt und keine neue Datensätze anlegt.

### **Duplizieren von Datensätzen:**

Das Duplizieren von Datensätzen ist eine wichtige Grundfunktionalität, die für eine effiziente Arbeitsweise wichtig wäre. Insbesondere beim Hochladen neuer, aber ähnlicher Datensätze, bei denen sich beispielsweise nur der Zeitraum, oder Kleinigkeiten bei der Bemessungsgrundlage geändert haben, würde diese Funktionalität eine große Zeitersparnis für Datenbereitsteller mit sich bringen. Dies sollte ohne allzu viel Programmieraufwand möglich sein.

*Low Hanging Fruits (Niedriger Aufwand, niedrig bis moderater Nutzen)*

### **Reminder-Funktion:**

Eine Reminder-Funktion soll primär dazu dienen den Bereitsteller des Datensatzes daran zu erinnern, dass dieser aktuell gehalten werden muss. Während sich der Datensatz an sich nicht verändert, können sich jedoch einige der Metadaten mit der Zeit ändern, was einer gewissen Pflege bedarf. Ein Beispiel dafür wäre, wenn sich der Ort, an dem die Ressourcen (Daten) auf der Website der jeweiligen Institution gespeichert sind, ändert und somit die URL aktualisiert werden muss.

### **Zurück-Button / In-View-Only Mode:**

Zum aktuellen Zeitpunkt ist es möglich, sich vorwärts als auch rückwärts durch den Wizard zu bewegen. Die Zurück-Funktion ist allerdings insofern bedenklich, als dass man nach einer abgeschlossenen Prüfung noch Felder editieren könnte die aber schon Gegenstand einer Prüfung waren. Hier würde es sich anbieten, Felder, oder ganze Module, nach Abschluss eines Arbeitsschrittes für die Bearbeitung zu sperren und nur noch im Rahmen eines View-Only-Modus einsehbar zu machen.



## Fahrplan zur Implementierung für Ministerien

GUIDO teil-automatisiert und vereinfacht den Prozess von der Datenauswahl bis zur Weitergabe der Metadaten an das Metadatenportal GovData. Es unterstützt Ministerien dabei ihrer Verpflichtung laut §12a des EGov-Gesetzes nachzukommen. Um GUIDO nutzen zu können, muss das von Tech4Germany entwickelte Sharepoint Webpart vom jeweiligen Ministerium integriert werden. Im Folgenden werden daher notwendige Schritte und Voraussetzungen, sowie wesentliche Inhalte der Implementierung beschrieben.

### Voraussetzungen zur Nutzung

GUIDO wurde in seiner aktuellen Form als sogenanntes Web Part für **SharePoint Online only (latest)** implementiert. Um das Tool in eine eigene OnPremise (auf eigenem Server installiert) SharePoint-Umgebung einbinden zu können, muss zunächst überprüft werden, welche SharePoint-Version vorliegt und was dies für eine Anpassung des GUIDO-Codes bedeutet.

Die Anbindung an das SharePoint Umfeld wird in GUIDO auf vier Arten verwendet:

- Zur Speicherung (Persistierung) der Prozessdefinitionen und der Bereitstellungsvorgänge als JSON Objekte wird **Lists** eingesetzt.
- Zur Speicherung der Daten, die im Bereitstellungsprozess hochgeladen werden (dies sind *nicht* die final veröffentlichten Daten) kommt **Documents** zum Einsatz, das an der jeweiligen SharePoint-Seite hängt, auf der das Web Part hinzugefügt wird. Dieser Ordner findet sich auch in **OneDrive**.
- Der Workflow zum Hochladen der Dateien per Email wird in **Power Automate** definiert. Die Anhänge der eingehende E-Mails werden in einem jeweils neu angelegtem Ordner in **SharePoint Documents** abgelegt und eine E-Mail mit einem Link zum neu angelegten Bereitstellungsvorgang geht zurück an den Sender.
- Zum Versenden von Benachrichtigungs-Emails wird **PnPjs** verwendet, dass den Umgang mit der SharePoint REST API vereinfacht

Für jede dieser SharePoint-Anbindungen ist zu klären ob, beziehungsweise unter welchen Bedingungen, dies in der jeweiligen SharePoint OnPremise-Installation

machbar ist. Falls die Möglichkeit nicht besteht, ließen sich relativ einfach Anbindungen an andere Datenbanken beziehungsweise an andere Speicherlösungen implementieren. Die Code-Struktur ist diesbezüglich bewusst so ausgelegt, dass alternative Anbindungen leicht einzubinden sind.

Ist kein SharePoint vorhanden, ließe sich GUIDO auch mit überschaubarem Aufwand aus der SharePoint Umgebung entfernen. Es handelt sich um eine React App, die bei entsprechender Umgestaltung einiger Stellen im Code die aktuell von einem SharePoint-Kontext ausgehen, auch unabhängig bestehen kann.

### *Aufwand der Implementierung*

Die vorher aufgezählten Voraussetzungen zur Nutzung zeigen aus, wieviel Aufwand jeweils für eine Implementierung notwendig wäre. Ohne die konkrete SharePoint-Version und die Beschränkungen hinsichtlich der Nutzerrechte zu kennen, ist es schwer hier eine pauschale Aussage zu treffen. Mit mindestens einer Woche Aufwand für einen Programmierer ist aber sicherlich zu rechnen. Es wäre gegebenenfalls lohnenswert einen SharePoint-Experten zu Rate zu ziehen, um Fragen bezüglich der Versionsanpassung und der Anbindung an das SharePoint-Ökosystem zu klären.

Der Code ist hier zu finden: <https://github.com/tech4germany/open-data-process-guide>

### *Prozessdefinition*

GUIDO ermöglicht es jedem Ministerium den Prozess der Datenbereitstellung anhand der eigenen Anforderungen und Wünschen zu definieren. Diese Prozessdefinition kann dabei nach bisherigem Stand an verschiedenen Punkten ansetzen.

Ist beispielsweise noch kein Prozess zur Bereitstellung definiert, kann die [Prozessbeschreibung zur Veröffentlichung von Daten](#) des CCODs als allgemeine Grundlage dienen und beliebig angepasst werden. Das CCOD kann darüber hinaus in weiteren Beratungsgesprächen helfen diesen Prozess weiter zu spezifizieren und eventuelle Unklarheiten zu klären.

Besteht Klarheit über den zu definierenden Prozess, kann dieser über zwei verschiedene Optionen in GUIDO integriert werden. Sofern im jeweiligen Ministerium Kompetenzen im Bereich **Business Process Modeling** bestehen, kann der Prozess in Form eines solchen erstellt werden und als *.bpmn-file* im Tool hochgeladen werden. Ein Beispiel für eine solche Prozessdefinition findet sich in Abbildung 1.

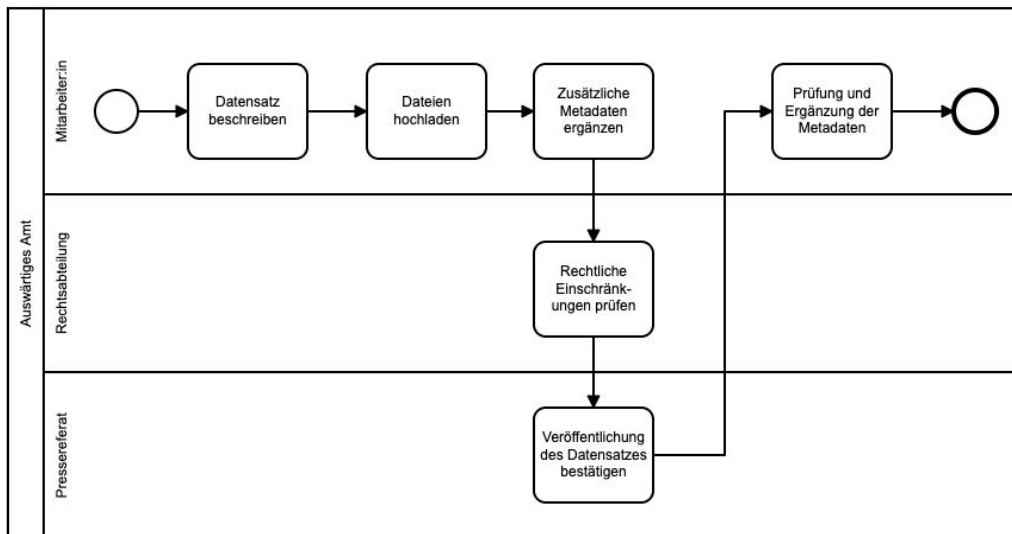


Abb 1: Prozessdefinition via Business Process Modeling and Notation (BPMN)

Falls dies nicht möglich oder gewünscht ist, kann der Prozess alternativ als *.json-file* definiert und hochgeladen werden. Programmierkenntnisse zur Erstellung sind dabei hilfreich aber nicht zwangsläufig notwendig. Eine Beschreibung der jeweiligen Module findet sich in der entsprechenden Konfigurationsdatei im [GitHub Repository](#). Die Namen der Arbeitsschritte im BPMN-Diagramm müssen mit den entsprechenden Modulen in der Konfigurationsdatei übereinstimmen, um korrekt eingelesen zu werden. Denkbar wäre hier noch mehr Definierungskompetenzen an BPMN abzugeben, so dass auch die auszufüllenden Felder pro Arbeitsschritt bereits in BPMN modelliert werden würden.

```

"processes": [
  {
    "id": "GuidoAA",
    "name": "Open Data Process Guide AA",
    "description": "The process of publishing Open Data in the Federal Foreign Office",
    "modules": [
      "describe-dataset",
      "data-upload",
      "supplement-metadata",
      "check-legal-constraints",
      "confirm-publication",
      "check-and-complete-metadata"
    ]
  }
]

```

Abb 2: Beispielhafte Definition des Prozess via json

### Übergabe der Daten an GovData

Diese Funktionalität haben wir im Rahmen des Fellowships nicht mehr umsetzen können.

### Wen fragen bei Fragen

Leider konnte bis zum Ende des Tech4Germany-Fellowships (Stand 19.10.2020) keine klare Verantwortlichkeit für GUIDO definiert werden. Bei allgemeinen Fragen und initialem Interesse am Thema Open Data empfehlen wir daher den Kontakt zum Kompetenzzentrum Open Data via [OpenData@bva.bund.de](mailto:OpenData@bva.bund.de)

Bei inhaltlichen Fragen zum Tool und der Nutzung empfehlen wir den Kontakt zu GovData via [info@govdata.de](mailto:info@govdata.de). Das Team von GovData hat während des Fellowships als Digitallotsen fungiert und stellt mit dem aktuellen Bereitstellungsformular den Vorreiter zu GUIDO bereit.

### Kommunikation des Tools

GUIDO ist nach der Implementierung für alle Mitarbeiter:innen innerhalb eines Ministeriums nutzbar. Das Tool ermöglicht Verwaltungsmitarbeiter:innen dabei ohne viel Vorwissen potentiell relevante Daten an das Tool zu schicken, beziehungsweise direkt im Tool hochzuladen, und somit unkompliziert den Prozess der

Datenbereitstellung zu starten, ohne die Verantwortung für alle notwendigen Schritte bis zur Veröffentlichung zu tragen.

Um auf diese Möglichkeit hinzuweisen und damit auch die Nutzungsrate von GUIDO zu steigern, empfehlen wir die Implementierung im Ministerien mit verschiedenen Kommunikationsmaßnahmen, wie beispielsweise einer gezielten Meldung im Intranet, über interne E-Mail Verteiler, oder auch dem Verfassen eines Hauserlasses zu begleiten.

## Strategische Überlegungen

Um mittel- und langfristig die Nachhaltigkeit und Nachnutzbarkeit von GUIDO zu sichern, bedarf es einer Einbettung der Ergebnisse in die weiteren Überlegungen und Entwicklungen im Bereich Open Data. Insbesondere das vom BMI geplante IT-Unterstützungstool verspricht eine Vielzahl von Vorteilen bei der Bereitstellung von Open Data für Ministerien. Im Folgenden wird daher kurz erläutert, inwiefern GUIDO und das geplante IT-Unterstützungstool sich hinsichtlich ihrer Funktionalitäten komplementieren und was es dabei gegebenenfalls zu beachten gilt.

### Integration in IT-Unterstützungstool

Während das CCOD insbesondere eine qualitative Beratung zum Thema Open Data bietet und Ministerien dabei unterstützt den Prozess der initialen Datenidentifizierung innerhalb eines Ministeriums zu starten, hilft GUIDO dabei den anschließenden Prozess der **Datenauswahl, Datenprüfung und Datenerfassung** zu vereinfachen und zu automatisieren.

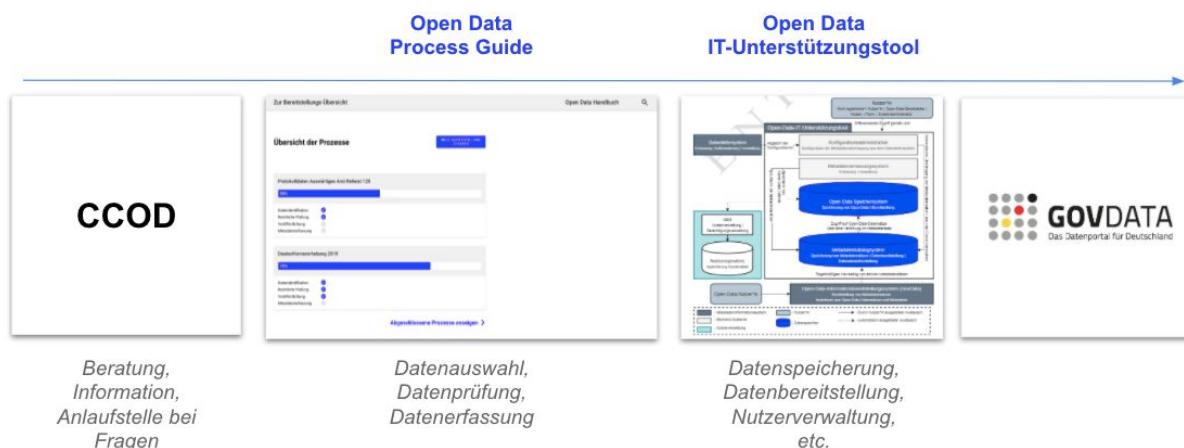


Abb 3: Abgrenzung verschiedener Open Data Angebote und Tools

Da sich das IT-Unterstützungstool aktuell (10/2020) noch in Planung befindet, basieren die folgenden Überlegungen auf den bereits bestehenden konzeptionellen Ausarbeitungen zum Tool. Wesentliche Inhalte und Vorteile sind die zentrale Datenspeicherung, Datenbereitstellung (zentraler Ort zur Weitergabe der Metadaten an GovData) und eine übergreifende Nutzerverwaltung für Ministerien.

Wie im [Fahrplan für Ministerien](#) dargelegt, kann vor der Fertigstellung des IT-Unterstützungstools die Weitergabe der Metadaten an GovData auch direkt über GUIDO mittels *.rdf-file* von einzelnen Ministerien an GovData erfolgen. Die Nutzerverwaltung und Speicherung der Daten erfolgt bei GUIDO über die Funktionalitäten von Microsoft Sharepoint.

Sobald das IT-Unterstützungstool entwickelt und implementiert ist, sollte die Übergabe der Metadaten von GUIDO an das IT-Unterstützungstool erfolgen, welches dann als zentraler Übergabepunkt der Metadaten aller Datensätze an GovData fungiert. Ferner können die Zwischenspeicherung noch nicht veröffentlichter Daten und das Nutzermanagement dann über die Datenspeicherung und das Nutzermanagement des IT-Unterstützungstools erfolgen.

Eine genaue Definition der Schnittstellen von GUIDO zum IT-Unterstützungstool sollte insbesondere bei der Entwicklung des IT-Unterstützungstools berücksichtigt werden, um eine technische Integration sicherzustellen.