Phase 2: Austausch, Vokabular, Veröffentlichung

1. Das passende Format für Haushaltsdaten

Zentral für die zweite Phase war eine Bestandsaufnahme mit den beteiligten Parteien. Dabei ging es konkret in ein moderiertes Gespräch mit Expert:innen aus der Haushaltsabteilung und den Menschen, die mit der technischen Umsetzung befasst sind.

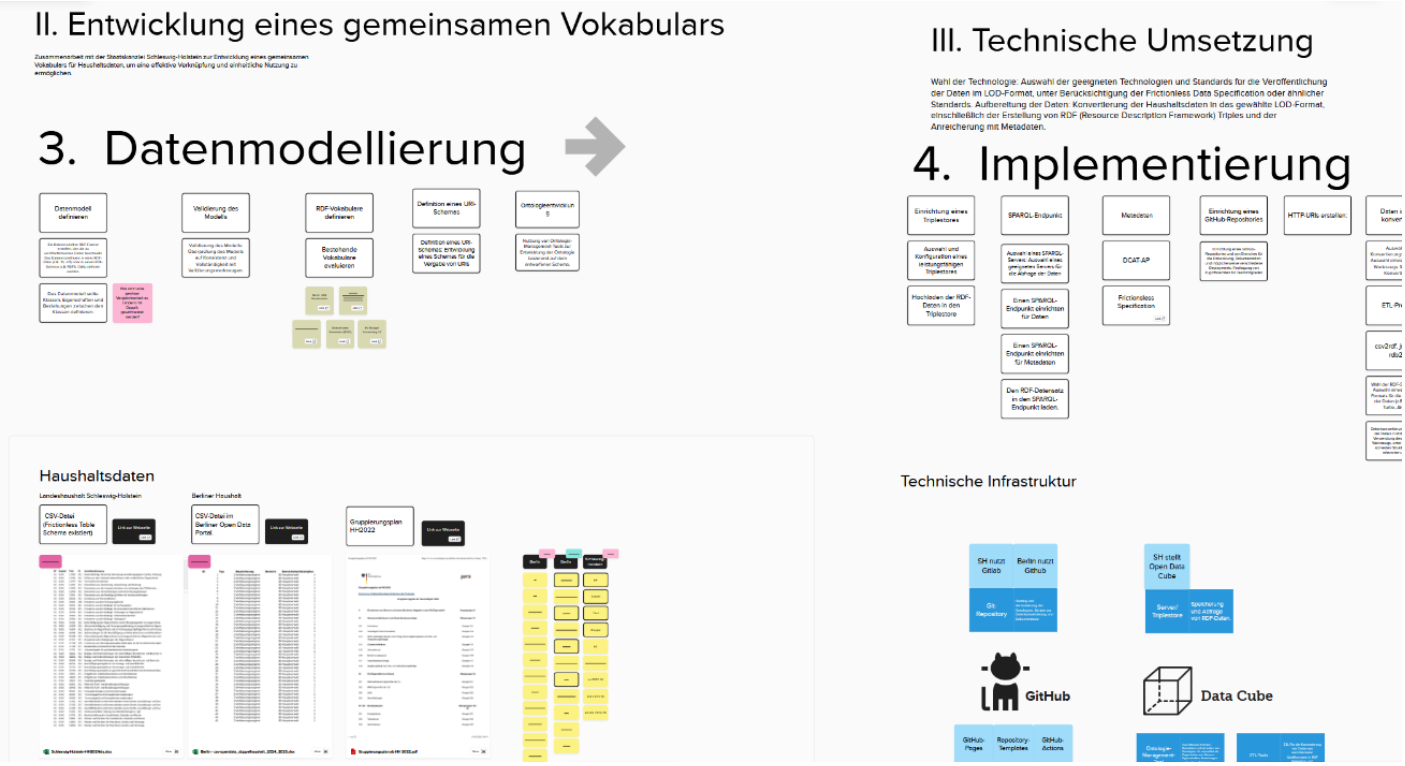


Bild: Ausschnitt aus dem gemeinsamen Arbeitsplan

Bei dem Gespräch wurden Unterschiede und Gemeinsamkeiten zwischen den Haushaltsdatensätzen der zwei Länder identifiziert und ein in Schleswig-Holstein entwickeltes Python-Skript für die Umwandlung von einer CSV zur RDF dem Team in Berlin zur Verfügung gestellt. Mit einigen Anpassungen konnten so die Haushaltsdaten beider Bundesländer in ein RDF-Format gegossen werden. Das Prinzip dahinter erklärt Jesper Zedlitz von der Staatskanzlei Schleswig-Holstein in einem Blogbeitrag. Für die Haushaltsdaten wurde sich an der existierenden Vorarbeit von Openspending - Linkedspending orientiert, die bereits vor mehr als zehn Jahren Haushaltsdaten in RDF Data Cubes modelliert haben.

2. Verlinkung ermöglichen - gemeinsames Vokabular für Haushaltsdaten entwickeln

Im Rahmen des gemeinsamen Gesprächs wurde festgestellt, dass für die Daten aus den Landeshaushalten bislang kein passendes Vokabular, das die eigentliche Verlinkung ermöglicht, existiert und deshalb selbst auf

Basis etwa der Ontologie des EU Haushalts und den Beispielen von Linkedspending erstellt werden muss. Bei der Datenmodellierung geht es um die Definition von Datenpunkten und ihrer Bedeutungen. In einem Vokabular wird mithilfe von Klassen und Eigenschaften festgelegt, welche Beziehungen zwischen diesen Datenpunkten möglich sind. Ziel ist es, ein konsistentes Schema für die Identifizierung von Daten zu entwickeln, das die Eindeutigkeit sicherstellt und gleichzeitig flexibel genug ist, um regionale und lokale Besonderheiten zu berücksichtigen.

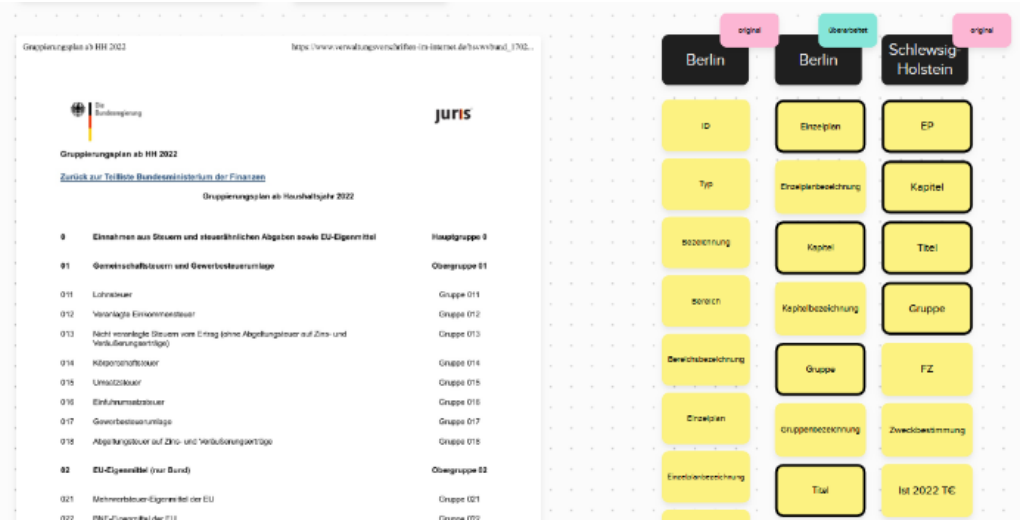


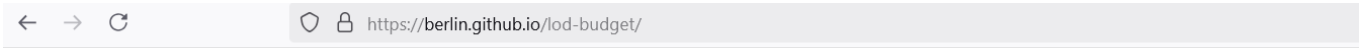
Bild: Beispiel für den Vergleich verschiedener Bereiche der Haushaltsdaten

Durch den Austausch zwischen den Haushaltsabteilungen der beiden Länder wurde festgelegt, worin der gemeinsame Teil zwischen den Haushaltsdatensätzen liegt oder wo gleiche Posten schlicht anders benannt sind. Für allgemeine Informationen sollte auf bereits etablierte Standards für Felder wie Jahreszahl etc. oder den Standard SDMX für statische Daten zurückgegriffen werden und so ein allgemeines Vokabular für Haushaltsdaten entstehen, das jeweils um landesspezifische Teile erweiterte werden kann.

3. Von RDF zu URIs

Die Definition von URIs bilden den Kern von Linked Open Data. Mittels einer URI lassen sich Daten online eindeutig identifizieren und wiederverwenden. Bei der Diskussion über die Verwendung von URIs zur Identifikation von Datensätzen zeigte sich der Ansatz, dass jede Verwaltungseinheit ihre eigenen Namensräume verwenden sollte, um Klarheit und Ordnung in der Datenorganisation zu gewährleisten. Der Namensraum gibt dabei Hinweis auf die Herkunft der Datenpunkte/des Vokabulars. Für die Haushaltsdaten hieß das konkret, dass die landesspezifischen Daten auf den Namespace des des jeweiligen Bundeslands verweisen, während für die übergreifenden Daten ein einheitlicher "Bundes"-Namespace genutzt wird. Mangels einer existierenden Basis aus dem Bund, wurde für das allgemeine Haushaltsvokabular die Plattform der Open Knowledge Foundation Deutschland genutzt.

Ein Weg, um aus existierenden RDFs diese Identifier zu kreieren und über Github zugänglich zu machen, wurde von BerlinOnline entwickelt. Mit Jinja-RDF können leicht statische Seiten erzeugt und angepasst werden. Da Berlin derzeit nicht über einen Triple Store verfügt, wurden die Haushaltsdaten selbst ebenfalls bei Github veröffentlicht.



LOD-Haushalt

Datensatz

Typ:	void:Dataset
Beschreibung:	Budgetary data for Berlin as RDF / Linked Open Data.
zuerst erstellt am:	2024-10-17
zuletzt geändert am:	2025-04-07
Veröffentlichende Stelle:	Senatsverwaltung für Finanzen
Homepage:	https://github.com/berlin/lod-budget
Beispielressource:	hhbe:6d3eeee8-2066-506e-a73a-4d23c0f759a2 (hh:Titel)
Beispielressource:	hhbe:e1c72463-dd22-5f4e-9119-89f43c83af74 (hh:Einnahmetitel)
Beispielressource:	hhbe:047106de-d059-54c7-b8ae-a2b132d4f1c8 (cube:Observation)
Beispielressource:	hhbe:2d61d0af-86a4-5458-8af5-8946d94e75dd (hh:Ausgabetitel)
Beispielressource:	hhbe:dd19576c-00df-53f6-9721-ab8c0dc0227a (cube:Observation)
Beispielressource:	hhbe:979e4bb7-0518-5a87-9693-d01f44c29099 (cube:Observation)
Genutztes Vokabular:	sdmxdim:
Genutztes Vokabular:	qudt:

Version #3
Erstellt: 19 Dezember 2024 10:04:52 von Admin
Zuletzt aktualisiert: 26 Juni 2025 09:18:12 von Admin