

Abschlussbericht



Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autor:innen.

Autor:innen:

Hannah Förster, Lukas Emele, Christian Winger, Ludwig Hülk, Christian Hofmann, Klara Reder, Mirjam Stappel, Martin Glauer

Verbundkoordination:

Fraunhofer-Institut für Energiewirtschaft und Energiesystemtechnik (IEE)

Projektpartner:

- 1) Reiner Lemoine Institut (RLI)
- 2) Öko-Institut e.V.
- 3) Otto-von-Guericke-Universität Magdeburg (OvGU)

März 2021

Gefördert durch:



Bundesministerium
für Wirtschaft
und Energie

aufgrund eines Beschlusses
des Deutschen Bundestages

Inhaltlicher Schlussbericht:

SzenarienDB – Datenbank für Klima- und Energieszenarien

Förderkennzeichen:

03ET4057A – Fraunhofer-Institut für Energiewirtschaft und
Energiesystemtechnik (IEE)

03ET4057B – Reiner Lemoine Institut gGmbH

03ET4057C – Öko-Institut e.V.

03ET4057D – Otto-von-Guericke-Universität Magdeburg

Projektträger:

Projektträger Jülich – Forschungszentrum Jülich GmbH (PtJ)

Projektzeitraum:

01.01.2018 – 31.03.2021

Projekt-Webseiten:

<https://www.iee.fraunhofer.de/de/projekte/suche/laufende/SzenarienDB.html>

<https://reiner-lemoine-institut.de/szenariendb/>

<https://www.oeko.de/forschung-beratung/projekte/pr-details/verbundvorhaben-szenariendb-datenbank-fuer-klima-und-energieszenarien>

<https://forschung-sachsen-anhalt.de/project/szenariendb-datenbank-klima-energieszenarien-22223>

Lizenz:

Sofern nicht anders angegeben, sind dieses Werk und seine Inhalte (Texte und Abbildungen) unter der Lizenz: Creative Commons - Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International (CC BY-SA 4.0 DE)



Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Wirtschaft und Energie unter den Förderkennzeichen 03ET4057A, 03ET4057B, 03ET4057C, 03ET4057D gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autor*innen.

Mitwirkende

Gedankt sei an dieser Stelle herzlich den zahlreichen Menschen, die die erfolgreiche Umsetzung des Projektes erst ermöglicht haben.

Seitens der am Projekt beteiligten Institute geht dieser Dank an die studentischen Hilfskräfte, PraktikantInnen, Bachelor- sowie Master-AbsolventInnen: Victoria Lister^{1, 2}, Izzet Kilikarslan², Jonas Huber², Lara Christmann², Ngoc Tu Phan², Johann Wagner¹⁰, Lukas Vogt¹⁰, Sebastian Cordes¹⁰, Anna Kleinau¹⁰, Simon Flügel¹⁰, Sebastian Cordes¹⁰, Sessa Sai Kiran Bhavaraju¹⁰, Sharanya Hunasamaranahalli¹⁰, Venkatesh Murugadas¹⁰, Dhruvin Suthar⁷, Martin Hertwig⁷.

Ein besonderer Dank geht an die Mitarbeiter*innen aller Projektpartner, die erfolgreich an der Realisierung des Projektes gearbeitet haben: Hannah Förster¹, Lukas Emele¹, Christian Winger¹, Ludwig Hülk², Christian Hofmann², Klara Reder³, Mirjam Stappel⁴, Carsten Pape⁵, Richard Schmitz⁶, Diana Böttger⁷, Martin Glauer⁸, Janna Hastings⁹, Meisam Booshehri¹⁰.

Des Weiteren geht unser Dank an die Community der *Open Energy Modelling Initiative (openmod)* und insbesondere an Max Muster für seinen aktiven Austausch und Treffen mit dem Projektteam. Auch möchten wir uns bei allen wissenschaftlichen MitarbeiterInnen und externen Studierenden, die auf Basis unserer Veröffentlichung ihre Arbeiten anfertigten und ihre Erkenntnisse dem Projektteam zurückgespielt haben, bedanken.

Für die Betreuung des Projektes geht unser Dank an den PtJ, der in allen administrativen Fragen und Entscheidungen ein verlässlicher Partner war. Schließlich danken wir dem BMWi für die Finanzierung des Projektes.

¹ Öko-Institut
² Reiner Lemoine Institut
³ Fraunhofer IEE
⁴ Fraunhofer IEE
⁵ Fraunhofer IEE
⁶ Fraunhofer IEE
⁷ Fraunhofer IEE
⁸ Otto-von-Guericke-Universität Magdeburg
⁹ Otto-von-Guericke-Universität Magdeburg
¹⁰ Otto-von-Guericke-Universität Magdeburg

Inhaltsverzeichnis

Mitwirkende	II
Abbildungsverzeichnis	VII
Tabellenverzeichnis	IX
0 Einleitung	3
0.1 Hintergrund	3
0.2 Motivation	3
0.3 Zielsetzung	4
0.4 Verantwortlichkeiten	5
1 Anforderungsanalyse für eine Szenariendatenbank	7
1.1 Erhebung der Anforderungen	7
1.2 Auswertung und Analyse der Anforderungen	14
1.3 Abgleich mit Funktionen der Open Energy Platform (OEP) und Erstellung des Lastenhefts	18
2 Ermittlung von Serveranforderungen für zukünftige Nutzung	23
2.1 Speicherbedarf abschätzen und Datenverkehr hochrechnen	23
2.2 Serverbetrieb	23
2.3 Hardware-Beschaffung- und Inbetriebnahme	24
3 Entwicklung des Konzeptes zur Weiterentwicklung der Datenbank	25
3.1 Ontologie der Energiemodellierungsdomäne	26
3.2 Vergleich existierender technischer Ansätze	29
3.3 Konzeptionierung	30
3.4 Softwaretechnische Modellierung im BPMN und UML	32
3.5 Kontinuierlicher Austausch mit der Community	34
4 Programmiertechnische Umsetzung des Konzeptes	37
4.1 Erweiterung des Open-Source-Konzeptes	37
4.2 Entwicklung Datenbankstruktur	38
4.3 Entwicklung und Optimierung der Datenbankschnittstellen	39
4.4 Erweiterung der Open Energy Platform (OEP)	40

5 Szenariendatenrecherche und exemplarische Befüllung der Datenbank	47
5.1 Recherche Szenarien	47
5.2 Prüfung der Verfügbarkeit der Daten für die Datenbank	50
5.3 Implikationen für Datenformate	51
5.4 Fragen der Datenlizenzierung	53
5.5 Integration verfügbarer Daten in die Datenbank	54
6 Entwicklung von Konzepten und Handlungsempfehlungen zur kontinuierlichen Aktualisierung/Erweiterung der Daten und zum Weiterbetrieb der Datenbank	59
7 Fazit	61
8 Ausblick	63
Literaturverzeichnis	65
A Anhang	i
A.1 Anhang nach Kapitel	i
A.2 Auswertung von Plattformen aus Nutzer*innensicht (AP 1.1)	vi
A.3 Vergleich existierender technischer Ansätze	vii
A.4 Erhebungsraster Szenariendaten (AP 5.1)	viii
A.5 Hochgeladene Tabellen (AP 5.5)	xiii
A.6 Konzept für den Weiterbetrieb (AP 6)	xxii

Abbildungsverzeichnis

1.1	MindMap zur Übersicht konkreter Personen, Gruppen und Institutionen als potenzielle Nutzer*innen der SzenarienDB.	9
1.2	Zuordnung von verschiedenen Merkmalen zu den Kategorien Unbedingt Erforderlich, Wichtig, Nice to have, Überflüssig und Keine Meinung.	13
1.3	Die bevorzugten Formate für Up- und Download sind .csv, .xlsx, API und Tabellenform.	14
1.4	Unterschiedliche Kriterien für eine Szenariodatenbank, geordnet nach Priorität der Klassen 1 bis 6.	15
1.5	Wortwolke abgeleitet aus den drei Freitextfragen: 'Folgende Features sind für mich bei der Suche von Daten im Internet/ Nutzung von Datenbanken außerdem wichtig bzw. unbedingt erforderlich', 'Welche Eigenschaften der OEP schätzen Sie bei der Nutzung?' und 'Welche Eigenschaften der OEP erschweren die Nutzung für Sie?'	16
1.6	Einstufung der Wichtigkeit von einer Vorschaufunktion als Tabelle von den beiden Nutzer*innengruppen APInists und Quantistits.	18
1.7	Der Prozentsatz der Anforderungen, die als sehr dringend (L), mittel dringend (M) und nicht dringend (S) eingestuft wurden.	20
1.8	Anforderungen in Prozent, aufgeteilt in zwölf Kategorien.	21
3.1	UML-Diagramm des Datenmodells des Modells appBBB	33
3.2	Business Process Model and Notation (BPMN) für den Review-Prozess beim Hochladens von Daten und Metadaten auf die OEP	33
4.1	Darstellung der Module der Open Energy Family	41
4.2	Das alte und das neue Logo der OEP (von links nach rechts)	42
4.3	Design der OEP vor dem Redesign	43
4.4	Design der OEP nach dem Redesign	43
4.5	Zusammenfassung aller erfassten Usability-Schwächen (10, S.62)	44

Tabellenverzeichnis

1.1	Identifizierte Nutzer*innengruppen für die SzenarienDB	8
1.2	Kriterien für Anforderungen	11
1.3	Auswahl von Ergebnissen aus der Onlineumfrage nach den Nutzergruppen APInist und Quantitist.	17
3.1	Öffentlichkeitswirksame Ereignisse in der Energiesystem-Community	35
3.2	Öffentlichkeitswirksame Ereignisse in der Energiesystem-Community Part II	36
A.1	Questions of the online survey on the topic: users expect/wish of a scenario database.	ii
A.1	Questions of the online survey on the topic: users expect/wish of a scenario database.	iii
A.1	Questions of the online survey on the topic: users expect/wish of a scenario database.	iv
A.2	Lesezeichentitel der Anhangstabelle	v

Abschlussbericht

SzenarienDB-Team

25. August 2021

0 Einleitung

0.1 Hintergrund

Dieser Bericht beinhaltet die Ergebnisse des Forschungsprojektes *Datenbank für Klima- und Energieszenarien (SzenarienDB)*. Der Bericht wurde gemeinsam durch die vier Verbundpartner das Fraunhofer-Institut für Energiewirtschaft und Energiesystemtechnik (IEE), das Reiner Lemoine Institut (RLI), das Öko-Institut e.V. (ÖI) und die Otto-von-Guericke-Universität Magdeburg (OVGU) erstellt. Für die Förderung des Verbundprojektes wird dem Bundesministerium für Wirtschaft und Energie (BMWi) ausdrücklich gedankt. Besonderer Dank gilt außerdem dem Projektträger Jülich (PtJ), namentlich Herrn Dr. Aleksandar Rakić, für die gute Betreuung des Projektes.

Der Abschlussbericht umfasst die Beschreibung der Weiterentwicklung der Open Energy Platform (OEP) mit Fokus auf der Erweiterung für Energieszenarien. Dabei liegen Schwerpunkte auf den Anforderungen der Nutzer*innen und der Entwicklung einer Ontologie für die Domäne der Energiesystemanalyse. Das Projekt folgt den Prinzipien von Open Science und kollaborativen Entwicklungen. Umgesetzt wird dies durch gemeinsamen Arbeiten auf GitHub (Open Source), dem Bereitstellen von offenen Daten in einer öffentlichen Datenbank (Open Data) und der Veröffentlichung von Artikeln, Präsentationen und Dokumentationen mit offenen Lizenzen (Open Access). Des Weiteren wurde ein Konzept zum Weiterbetrieb der Datenbank entwickelt.

0.2 Motivation

Bei dem komplexen Prozess der Transformation des Energieversorgungssystems spielen eine Reihe von Modellen, darunter Energiesystemmodelle und Modellverbände eine zunehmend wichtige Rolle, um die Auswirkungen möglicher zukünftiger Entwicklungen zu ermitteln, Effekte von Klimaschutzpolitiken zu bewerten, sowie um optimale bzw. geeignete Entwicklungspfade zu identifizieren.

Szenarien, die einen möglichen Entwicklungspfad bei der Umgestaltung des Energieversorgungssystems abbilden, werden zur Parametrisierung für alle Energiesystemmodellierungen benötigt. Die Wahl des Szenarios entscheidet in großem Maße über die Ergebnisse und beeinflusst erheblich die Handlungsempfehlungen, die aus Energiesystemmodellierungsstudien abgeleitet werden (7). Die Vielzahl unterschiedlicher Szenarien, die als Eingangsdaten für Energiesystemmodellierungsstudien dienen, führten zusammen mit der Verwendung unterschiedlicher Modelle und Modellierungsansätze in der Vergangenheit dazu, dass die Ergebnisse von Studien nur schwierig bewertet und miteinander verglichen werden können. In verschiedenen Projekten wurden bereits für Modellvergleiche Eingangsdaten und Annahmen abgeglichen, um die Modellverhalten zu vergleichen (MEX (4) sowie RegMex(8)), die Szenariendaten wurden jedoch nicht veröffentlicht.

Ein Beispiel, wo versucht wurde, Ergebnisse verschiedener Studien systematisch zu vergleichen und zu bewerten ist die im Auftrag des BMWi von Fraunhofer IWES und Fraunhofer Umsicht durchgeführte „Metastudie Energiespeicher“ (5). Die Studie zeigt deutlich, wie schwierig eine solche Bewertung ist, da sowohl Szenarien als auch eingesetzte Modelle stark variieren und schließlich zu teilweise recht unterschiedlichen Ergebnissen führen. Ein weiteres Problem stellt die mangelnde Transparenz dar. Nur sehr selten werden in Studien die detaillierten Szenarioannahmen veröffentlicht und mit geeigneten offenen Lizenzen versehen.

0.3 Zielsetzung

Aus den genannten Gründen ist es das Ziel des Projektes „SzenarienDB“ eine Energieszenariendatenbank als Erweiterung der Open Energy Platform (OEP) mit zusätzlichen Funktionalitäten zu erstellen. Es wird eine offene Community-Datenbank für Energieszenarien und Energiesystemdaten entwickelt, die als Grundlage für Transparenz und Vergleichbarkeit in der Energieforschung genutzt werden kann.

Die Ziele für die Energieszenariendatenbank sind:

- Zugang durch ein benutzer*innenfreundliches Webinterface
- Zugriff durch eine API unabhängig des Betriebssystems oder der Programmiersprache
- Versionierung der Daten, eingeschlossen alter Ergebnisse, Korrekturen von Fehlern, Hinzufügen von Szenarien
- Vorbildfunktion und Nutzbarkeit für ähnliche Projekte in anderen Disziplinen oder Regionen

- Start einer breiten Diskussion über Standards für den Datenaustausch und Modellbeschreibungen für Szenarien
- Entwicklung einer Ontologie für die Energiesystemanalyse

Neu an unserer Datenbank, im Vergleich zu anderen Energiedatenbanken, ist:

- Die erste offene Community-Datenbank für Energieszenarien
- Reduzierung von Aufwand in der Energiemodellierung durch den einfachen Zugang zu bestehenden Energieszenarien
- Eine gemeinsame Datenbank für Eingangs- und Ausgangsdaten

0.4 Verantwortlichkeiten

Die Verantwortlichkeiten zwischen den Projektpartnern ist wie folgt aufgeteilt:

- **Fraunhofer IEE** Projektleitung, Lastenhefterstellung, Konzeptionierung, Einbringen von Domänenwissen, Szenariupload, Konzepte für den Weiterbetrieb der Datenbank,
- **Öko-Institut e.V.** Szenarienauswahl und -Upload, Einbringen von Domänenwissen.
- **Reiner Lemoine Institut** Pflichtenhefterstellung, Einbringen von Domänenwissen, Szenariupload, Konzeptionierung für die Erweiterung der Datenbank.
- **Otto-von-Guericke-Universität Magdeburg** Programmtechnische Umsetzung des Konzeptes, Ermittlung der Serveranforderungen.

1 Anforderungsanalyse für eine Szenariendatenbank

Im Arbeitspaket 1 wurden die Grundlagen für den Aufbau der Szenariendatenbank gelegt. Zunächst wurden die Anforderungen an die Datenbank erhoben. Hier wurde in Nutzer*innenanforderungen und strukturelle Anforderungen unterschieden. Die Anforderungen wurden analysiert und schließlich in einem Lastenheft dokumentiert.

Die grundlegenden Fragen wurden geklärt:

- Welche Zielgruppen stellen potenzielle Nutzer*innen für die SzenarienDB dar?
- Welche Anforderungen haben die verschiedenen Nutzer*innengruppen?
- Welche strukturellen Anforderungen lassen sich ableiten?
- Welche Funktionen und Module müssen weiterentwickelt werden?

1.1 Erhebung der Anforderungen

In diesem Abschnitt wird die Erhebung der Nutzer*innenanforderungen aus AP 1.1 und der Strukturanforderungen aus AP 1.2 gemeinsam beschrieben, da sie aus den gleichen Quellen abgeleitet wurden und einander zum Teil entsprechen. Zunächst wurden verschiedene Nutzer*innengruppen identifiziert. Dann wurden mit verschiedenen Methoden die Anforderungen erhoben, unter Berücksichtigung der verschiedenen Gruppen.

Definition der Nutzer*innen/-gruppen Zunächst haben wir drei mögliche Nutzer*innengruppen identifiziert, die sich durch unterschiedliche Charakteristika auszeichnen, und unterschiedliche Anforderungen an eine Plattform mit Szenarien haben. Besonders deren Programmieraffinität nutzen wir als Merkmal. Die entstandenen Nutzer*innengruppen bezeichneten wir als „API-nists“, „Quantitists“ und „Qualitists“. Die verschiedenen Merkmale sind in der folgenden Tabelle 1.1 aufgeführt. Aus diesen drei theoretischen Gruppen leiteten wir im nächsten Schritt exemplarisch konkrete Personen und Institutionen ab, die bei der Ermittlung der Nutzer*innenanforderungen hilfreich sein konnten - siehe Abbildung 1.1.

Tabelle 1.1: Identifizierte Nutzer*innengruppen für die Szenariendatenbank

Charakteristika	„API-nists“	„Quantitists“	„Qualitists“
Bevorzugter Datenzugriff und Datenmerkmale	API, große Datenmengen, detailliert, numerisch	mittelgroße Datenmengen, Tabellen	kleine Datenmengen, leicht zugänglich, aggregiert, übersichtlich
Bevorzugte Formate	Standardisiert, maschinenlesbar	Tabellen	Text, Tabellen
Bevorzugte Zusätze	Metadaten, Ontologie	Datenlexika, Metadaten	fertige Visualisierungen
Beispiele	Programmierer*innen, Informatiker*innen	Naturwissenschaftler*innen, Ökonom*innen	Sozialwissenschaftler*innen
Personen/ Institutionen	Projektpartner*innen, Modellierer*innen, openMod	Modellierer*innen, Strommarktgruppe, Studierende, JRC	Auftraggeber*innen, Desk Officers EU-Kommission, NGOs, BMWi, Beirat Systemanalyse

In einem nächsten Schritt werteten wir 15 Datenplattformdienste aus dem Blickwinkel dieser drei Nutzer*innengruppen und nach verschiedenen Merkmalen aus. Die berücksichtigten Merkmale waren dabei:

- data: license, temporal coverage, temporal resolution, spatial coverage, spatial resolution, coverage / resolution of other dimensions
- data search: search / index, glossary, data preview, on-site visualisation, data description / metadata
- data retrieval: API availability, download formats, download selection / filter, download restrictions
- user friendliness: clear user interface, accessibility for non-programmers, does main page hold relevant information (license, user feedback),

Die berücksichtigten Plattformdienste waren: UN Comtrade, World Bank Open Data, Eurostat, European Union Open Data, Open Street Map, Testreferenzjahre, Climate Data Centre, EUMESAT, FAOSTAT, Open Energy Platform, OpenEI, Energy Data, reegle, European Data Portal und UNdata. Die Auswertung erfolgte nach den oben genannten Merkmalen, für jede Nutzer*innengruppe und gewichtet

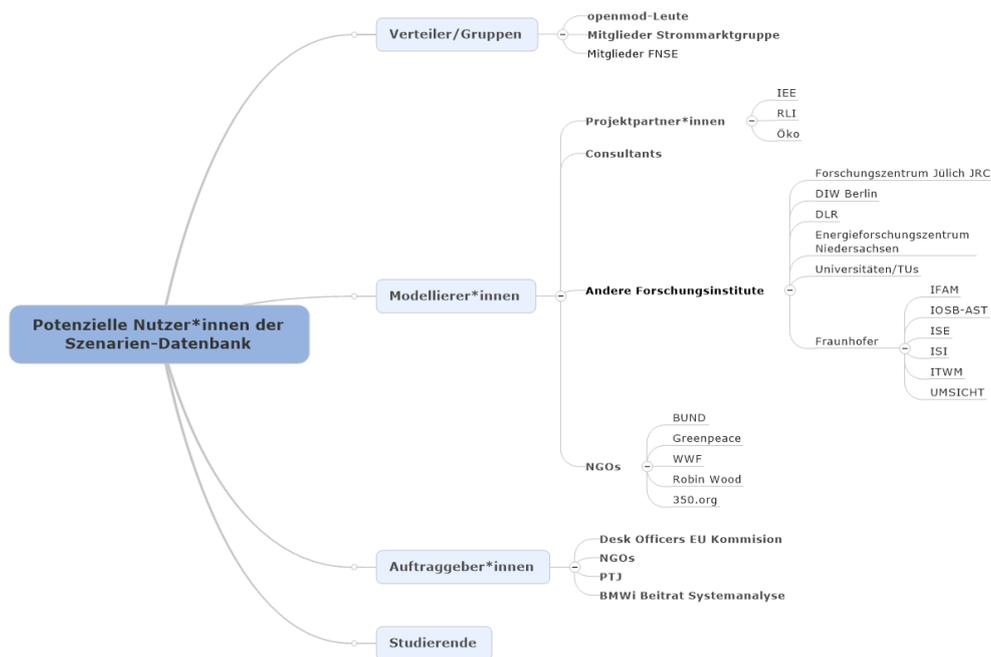


Abbildung 1.1: MindMap zur Übersicht konkreter Personen, Gruppen und Institutionen als potenzielle Nutzer*innen der SzenarienDB.

mit der von uns angenommenen Wichtigkeit des Merkmales für die jeweilige Nutzer*innengruppe. Dabei gingen wir z.B. davon aus, dass APInists weniger Wert auf ein klares Interface legen als Qualitists und erstere relativ mehr Wert auf das Vorhandensein einer API Schnittstelle. Die Bewertungsmatrix mit der Gewichtung der Merkmale je Nutzer*innengruppe, sowie der Ergebnisse befindet sich im Appendix unter A.2.

Die beiden am besten bewerteten Plattformdienste – FAOSTAT und Open Street Map – unterzogen wir dann in AP 3.2 einer technischen Analyse, deren Ergebnisse Input in das Lastenheft zur Verfügung stellten.

Methodisches Vorgehen Um die Anforderungen zu ermitteln wurden drei Wege eingeschlagen. Der erste Weg war eine Online-Umfrage bei potenziellen Nutzer*innen. Die unter Abschnitt ?? beschriebenen Zielgruppen konnten hierfür über verschiedene E-Mail-Verteiler und Plattformen mit Fokus auf Energiesystem-Themen weitgehend erreicht und zu einer Teilnahme an der Umfrage eingeladen werden: Strommarktgruppe, Openmod-Verteiler und Forschungsnetzwerk Energiesystemanalyse (FNE), der Verteiler der Richtlinien-Gruppe zu Energieszenarien sowie interne Verteiler am PtJ, BMWi, Öko-Institut, RLI und IEE. Zweitens wurden Expert*innen-Workshops durchgeführt, in denen direkt über Nutzer*innen- und Strukturanforderungen diskutiert wurde und aus denen ebenfalls Anforderungen abgeleitet werden

konnten. Und drittens wurden die bisher im GitHub gesammelten „Issues“ gesichtet und daraus Anforderungen abgeleitet. Die Anforderungsanalyse wurde in (12) veröffentlicht und in einem Poster auf der ESS Conference 2018 in Karlsruhe vorgestellt.¹

Nutzer*innen- und Strukturanforderungen können mit verschiedenen Methoden bestimmt werden, dazu gehören Interviews, Vergleich mit anderen Systemen, Beobachtung der Benutzer*innen während der Anwendung etc. (13). Hier wurden Interviews in Form einer Online-Umfrage als geeignete Methode ausgewählt. Diese Methode wurde aufgrund der Erreichbarkeit der Zielgruppe, der Zeitersparnis und der niedrigen Kosten gewählt (6, 15). Die Hauptfrage war: „Was sind die Nutzer*innen- und Strukturanforderungen an eine Open-Source-Datenbank die Energieszenarien enthält“. Die Online-Umfrage bestand aus zwei Teilen, im ersten Teil wurden Fragen zur tägliche Arbeit mit Szenarien gestellt. Im zweiten Teil richteten sich die Fragen an die Merkmale und Kriterien, die eine Szenariendatenbank haben oder erfüllen sollte. Sie wurde auf deutsch und englisch gestellt. Eine vollständige Liste aller Fragen ist im Anhang A.1 zu finden. Der Zeitraum der berücksichtigten Antworten liegt zwischen dem 12. Juni und dem 27. August 2018.

Des weiteren fanden einige Meetings statt, in denen Expert*innen aus dem Projekt jeweils ein Thema diskutierten. Diese Themen waren:

- Welche Metadaten sollen aufgenommen werden?
- Wie sollen Referenzen gesetzt werden können?
- Wie kann man hochgeladene Daten überprüfen - Reviewprozess?
- Anforderungen an Tutorials, z.B. des oedialect?

Aus den Antworten der Online-Umfrage und den Freitextfeldern sowie aus den Expert*innen-Meetings leiteten wir Anforderungen ab und trugen diese in GitHub zusammen. Sie sind einheitlich in Englisch verfasst. Zudem glichen wir existierende GitHub-Issues mit diesen Anforderungen ab, siehe Abschnitt 1.2, und ergänzten gegebenenfalls².

Es wurde angestrebt, den Anforderungen die Struktur von ‘User Stories’ zu geben: ‘Als *<Rolle >*möchte ich *<Ziel/Wunsch>*[um *Nutzen*]]’ (3). Hier ein Beispiel einer ausformulierten Anforderung: ‘As a user I want to search for texts to find single entries in a dataset.’ Darüber hinaus sollten alle Anforderungen so formuliert sein, dass sie die Kriterien aus Tabelle 1.2 erfüllen (2, 11).

¹https://www.iee.fraunhofer.de/content/dam/iee/energiesystemtechnik/de/Dokumente/Poster/2018/2018_Poster_SzenarienDB_v5_Lizenz.pdf

²https://github.com/OpenEnergyPlatform/oeplatform/issues?q=is%3Aopen+is%3Aissue+label%3Aspecification_sheet

Eine andere gängige Methode zur Formulierung von Anforderungen ist die INVEST-Methode (14). Hier lauten die Kriterien für die User Stories ‘unabhängig’, ‘verhandelbar’, ‘für die Kunden wertvoll’, ‘schätzbar’, ‘klein’ und ‘testbar’. Alle diese Kriterien sind in den im Projekt angewandten Kriterien enthalten, außer ‘verhandelbar’. ‘Verhandelbar’ wird jedoch erfüllt durch die Veröffentlichung der Anforderungen im GitHub-Repository des OEP, wodurch öffentlich an den Anforderungen mitgearbeitet werden konnte.

Tabelle 1.2: Kriterien für Anforderungen

Kriterium	Beschreibung
Priorität	Alle Anforderungen werden anhand einer Priorität bewertet.
Vollständigkeit	Vollständigkeit erfordert, dass alle Aspekte der Anforderung ohne implizite Annahmen formuliert werden.
Dokumentation/ Nachvollziehbarkeit	Jede Anforderungen muss über eine Dokumentation darüber verfügen, woher die Anforderung stammt und wie die Entwicklung bis zur endgültigen Formulierung verläuft um damit Nachvollziehbar zu sein.
Korrektheit	Die Anforderungen sollen korrekt sein und von den Beteiligten als notwendig akzeptiert werden.
Verständlichkeit	Die Anforderungen sind als nicht-technischer verständlicher Text formuliert, so dass alle Beteiligten, vom Stakeholder bis zum Entwickler, verstehen, was gemeint ist.
Konsistenz	Die Anforderungen sollen konsistent sein. Das bedeutet, dass keine konkurrierenden Interessen zwischen den Anforderungen bestehen.
Überprüfbarkeit	Alle Anforderungen können getestet werden, um sicherzustellen, dass sie funktionstüchtig sind.
Eindeutigkeit	Die Anforderungen sind eindeutig, das heißt das sie jeweils nur ein Problem beschreiben.
Aktion	Die Anforderungen sollten eine Aktion beschreiben, die ein Nutzer durchführen kann.

Ergebnisse der Onlineumfrage

Die Online-Umfrage wurde von 177 Personen gestartet und von 101 Personen vollständig beendet. Im Folgenden werden nur die vollständig ausgefüllten Umfragen berücksichtigt. Der Großteil der Personen ist in Forschungseinrichtungen tätig (71%).

Viele TeilnehmerInnen (69%) sind sowohl an der Szenarien-Generierung als auch an der Nutzung fremder Szenarien beteiligt. Nur 6 Personen arbeiten überhaupt nicht mit Energieszenarien. Etwa 56% nutzen häufig externe Datenbanken wie Eurostat, OpenStreetMap und andere, nur 11% verwenden überhaupt keine Datenbanken.

Die Umfrage zeigte, dass es unter den Energiemodellierer*innen ein großes Interesse an dem Thema Open Source/Open Data gibt. Die Teilnehmer*innen sind bereit, Energieszenarien aus einer solchen Datenbank zu nutzen (96%) und dort eigene Szenarien zu veröffentlichen (92%). Voraussetzung für die Veröffentlichung von Szenarien ist für Viele, dass die Finanzierung gegeben ist. Hindernisse liegen in der Lizenzierung oder der kommerziellen Nutzung von Szenarien. Die Teilnehmer*innen wurden nach ihrer Bereitschaft gefragt, eine API zwischen SzenarienDB und ihren Modellen zu realisieren. Die Mehrheit (53%) war bereit, dies unter bestimmten Bedingungen zu tun, 26% auch ohne Bedingungen.

Die Mehrheit der Teilnehmer*innen (52%) benötigt hochauflösende Daten (z.B. stündliche Zeitreihen für ein Jahr, räumliche Auflösung im Maßstab von Kilometern). Nur 19% verwenden Daten mit geringem Detaillierungsgrad, wie z.B. aggregierte Werte für Länder oder Jahre.

Darüber hinaus ist die Qualität (52%) für die Teilnehmer*innen am wichtigsten, gefolgt von Quantität (25%) und Benutzer*innen-Freundlichkeit (23%).

Die Teilnehmer*innen wurden gebeten, sechs Merkmalen eine unterschiedliche Wichtigkeitsstufe zuzuweisen. Abbildung 1.2 zeigt die Ergebnisse in absteigender Reihenfolge: 'Filtern von Daten', 'Beschreibung der Metadaten', 'Textsuche', 'Glossar/Onologie', 'Vorschau der Daten' und 'Ad-hoc-Visualisierung'. Die Möglichkeit, Daten zu filtern, wurde mit Abstand am häufigsten (70%) als unerlässlich eingestuft. Für die Beschreibung der Metadaten, die Textsuche, das Glossar/die Ontologie und die Vorschau der Daten wurden die Kategorien wichtig und unbedingt erforderlich ähnlich häufig genannt, wobei unbedingt erforderlich noch etwas häufiger genannt wurde. Die Funktion Ad-hoc-Visualisierung wurde von den meisten Teilnehmer*innen in die Kategorie "nice to have" eingestuft (60%). Nur sehr Wenige ordneten einem der Features die Kategorie "überflüssig" ($\leq 7\%$) oder "keine Meinung" ($< 4\%$) zu.

Die bevorzugten Formate für Up- und Download sind .csv, .xlsx, API und Tabellenform, wie in Abbildung 1.3 dargestellt. Die Teilnehmer*innen hatten die Möglichkeit, mehrere Formate auszuwählen.

In einer weiteren Frage sollten die Teilnehmer*innen die Klassen eins (ist mir am wichtigsten) bis sechs (ist mir am wenigsten wichtig) verschiedenen Kriterien zuordnen (Abbildung 1.4). Eine 'Quellen-/Literaturangaben zu allen Datensätzen' wurde am häufigsten (56%) für die Klasse 1 (höchste) ausgewählt. Die Verteilung über die

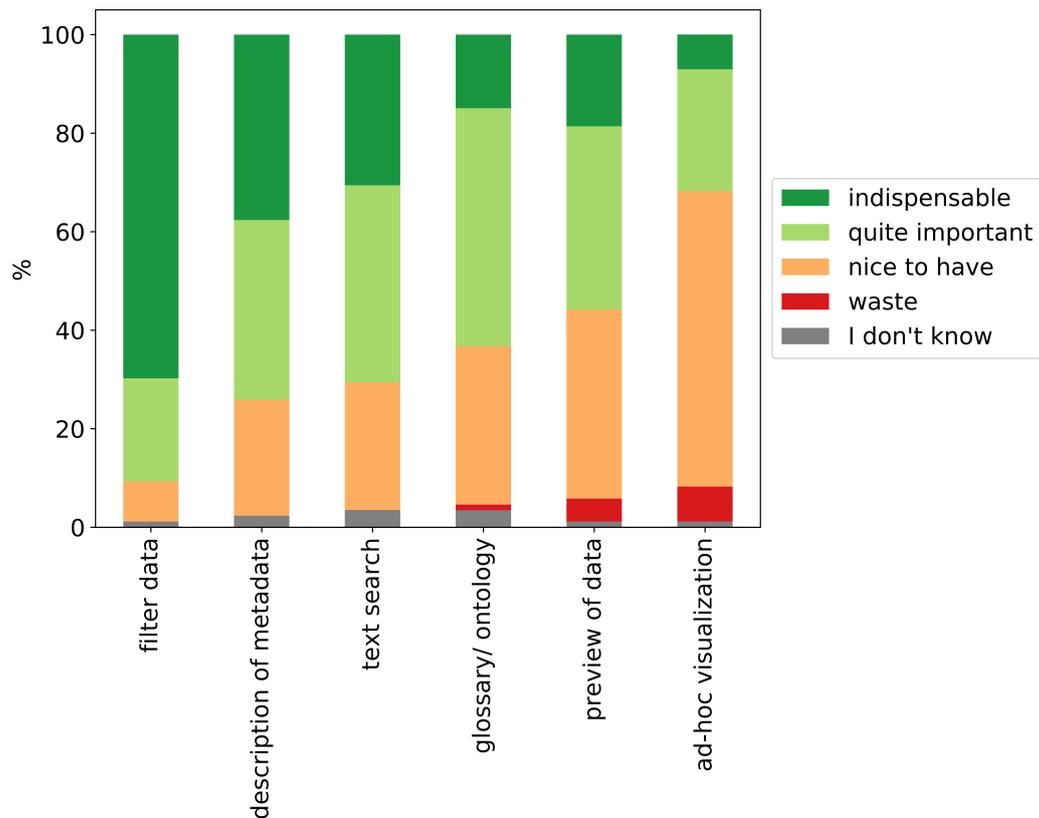


Abbildung 1.2: Zuordnung von verschiedenen Merkmalen zu den Kategorien Unbedingt Erforderlich, Wichtig, Nice to have, Überflüssig und Keine Meinung.

verbleibenden Klassen ist absteigend mit absteigender Wichtigkeit. Das zweit häufigste Kriterium in der Klasse 1 ist die 'Qualitätskontrolle upgeladeter Szenariendaten durch die Betreiber', aber der höchste Prozentsatz (40%) für dieses Kriterium wird für die Klasse 2 gewählt. Die Verteilung auf die niedrigeren Klassen ist ähnlich, jeweils zwischen 7-12%. Die Kriterien 'Einfacher, intuitiver Upload von eigenen Szenarien' und 'Schnelligkeit' haben eine ähnliche Verteilung. Die Teilnehmer wählten am häufigsten eine Klasse 3-4 (zwischen 19-34%) und weniger oft eine hohe Klasse 1-2 oder eine niedrige Klasse 5-6. Die am wenigsten wichtigen Kriterien sind die 'Möglichkeit, Daten direkt in der Datenbank zu prozessieren' und 'Konvertierung von Einheiten' (27% bzw. 33% Klasse 6).

In der Online-Umfrage enthielten drei Fragen Freitext: 'Folgende Features sind für mich bei der Suche von Daten im Internet/ Nutzung von Datenbanken außerdem wichtig bzw. unbedingt erforderlich', 'Welche Eigenschaften der OEP schätzen Sie bei der Nutzung?' und 'Welche Eigenschaften der OEP erschweren die Nutzung für Sie?' Alle Antworten wurden in einer Wortwolke in der Abbildung 1.5 zusammengefasst. Wörter wie 'z.B.' und 'to' (englisch 'zu') wurden entfernt. Außerdem wurde die

Abbildung 1.3: Die bevorzugten Formate für Up- und Download sind .csv, .xlsx, API und Tabellenform.

Konjugation deutscher Wörter vereinheitlicht (z.B. 'keine' und 'kein'). Die daraus resultierende Worldcloud gibt einen guten Überblick über die Themen, mit denen sich die Teilnehmer am meisten beschäftigen. Die Online-Umfrage wurde hauptsächlich auf deutsch, aber auch in englischer Sprache beantwortet. Die gleichen deutschen und englischen Wörter wurden nicht vereinheitlicht, so dass die Abbildung 1.5 am stärksten ausgeprägt 'Daten' zeigt, aber auch die englische Version 'Data' zu finden ist.

1.2 Auswertung und Analyse der Anforderungen

Es wurden die Nutzer*innengruppen APInists (Programmierer, Modellierer), Quantitists (Naturwissenschaftler, Ingenieure, Ökonomen) und Qualitists (Sozialwissenschaftler) identifiziert. In der Onlineumfrage hatten die Teilnehmer die Möglichkeit mehrere Antworten bezogen auf ihre Tätigkeit anzugeben. Eine Filterung auf die Nutzer*innengruppen hat ergeben, dass 69 APInists, 76 Quantitists und 2 Qualitists an der Umfrage teilgenommen haben. Wobei zu beachten ist, dass durch die Mehrfachnennung keine eindeutige Unterscheidung vorhanden ist, da es bei der Umfrage nur 101 Teilnehmer gab und dementsprechend Mehrfachnennung vorgekommen sind. Es haben nur zwei Qualitists an der Umfrage teilgenommen, woraus folgt, dass dies entweder keine Nutzer*innengruppe von relevanter Größe ist, oder diese Nutzer*innengruppe durch die Umfrage nicht erfasst wurde. Da der Umfragebogen an einschlägige E-Mail Verteiler verschickt wurde, sind die Qualitists vermutlich nicht Teil dieser Verteiler. Aufgrund der geringen Stichprobe von Qualitists werden diese in der folgenden Auswertung nicht mit betrachtet.

Die APInists und Quantitists haben die Fragen weitgehend gleich beantwortet, so dass sich für das Lastenheft keine gesonderten Anforderungen bezogen auf die Nut-

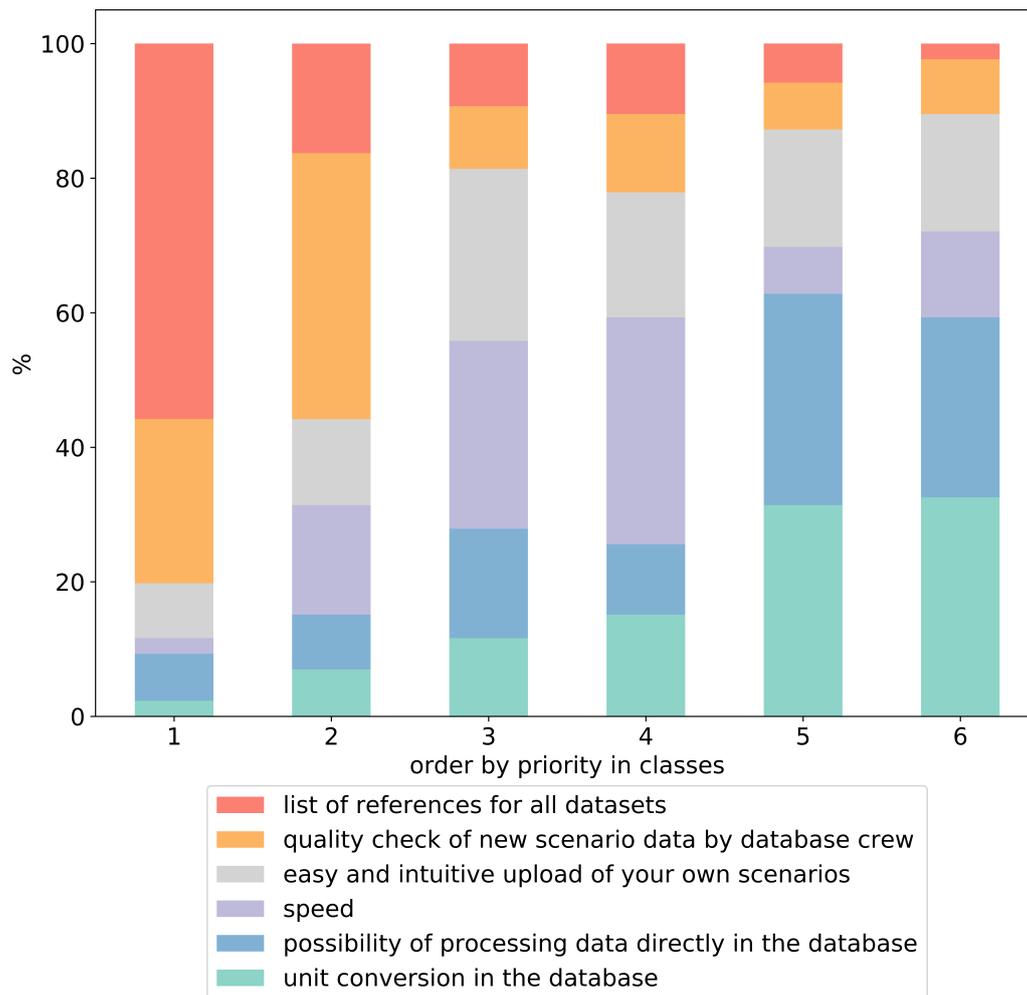


Abbildung 1.4: Unterschiedliche Kriterien für eine Szenariodatenbank, geordnet nach Priorität der Klassen 1 bis 6.

zer*innengruppen erstellen lassen. Da es sich durch die Mehrfachnennung z.T. um die gleichen Personen handelt, ist es nicht überraschend, dass die Antworten sehr ähnlich sind. Für zukünftige Umfragen wäre es eine Überlegung, dass die Befragten nur eine Wahlmöglichkeit haben. Diese Wahlmöglichkeit könnte als Schwerpunkttätigkeit beschrieben werden. Damit hätten die Nutzer*innengruppen anschließend möglicherweise unterschiedlichere Ausprägungen. Die APInists (78%) arbeiten etwas mehr an Forschungsinstituten als die Quantitists (71%). Die Quantitists arbeiten dafür zu gleichen Teilen mehr in den Bereichen Consulting, Öffentliche Auftraggeber und Netzbetreiber. APInists sind laut Nutzer*innengruppen-Definition technisch affiner und mehr an der Erstellung von Szenarien beteiligt, was sich auch in den Ergebnissen der Umfrage zeigt, wie in Tabelle 1.3 zu sehen ist. Auch können sich weniger Quantitists vorstellen eine Schnittstelle einzurichten was vermutlich durch den komplexen



Abbildung 1.5: Wortwolke abgeleitet aus den drei Freitextfragen: 'Folgende Features sind für mich bei der Suche von Daten im Internet/ Nutzung von Datenbanken außerdem wichtig bzw. unbedingt erforderlich', 'Welche Eigenschaften der OEP schätzen Sie bei der Nutzung?' und 'Welche Eigenschaften der OEP erschweren die Nutzung für Sie?'

Aufwand zu erklären ist. Dass den Quantitists eine leichtere Bedienung wichtiger ist sieht man auch an der Priorisierung der Features. Den Quantitists ist eine Vorschaufunktion wichtiger als den APInists. Den APInists hingegen sind eine Beschreibung der Metadaten als Text , Quellen/- Literaturangaben und Glossar/Ontologie wichtiger als den Quantitists. Es ist jedoch zu beachten, dass diese Unterschiede nur für einen Teil der jeweiligen Nutzer*innengruppen gilt, wie man exemplarisch an Abbildung 1.6 für die Vorschaufunktion sieht.

Alle anderen Fragen wurden von den beiden Nutzer*innengruppen derart ähnlich beantwortet, dass keine signifikanten Unterschiede erkennbar waren. Es wurde noch untersucht, ob es eine Nutzer*innengruppe gibt, die keine Szenarien einspielen würde. Allerdings haben das nur sieben Personen angegeben, die aus den unterschiedlichsten Bereichen kamen, so dass hieraus keine spezielle Nutzer*innengruppe abgeleitet werden konnte. Genauso konnte aus den elf Personen die angegeben haben keine Schnittstelle verwenden zu wollen keine Nutzer*innengruppe abgeleitet werden. Die

Tabelle 1.3: Auswahl von Ergebnissen aus der Onlineumfrage nach den Nutzergruppen APInist und Quantitist.

Aussage	APInists %	Quantitists %
Erstellen Szenarien und nutzen die Szenarien anderer	77	66
Nutzen nur die Szenarien anderer	11	23
Sind bereit Szenarien in SzenarioDB zu veröffentlichen	53	58
Sind bereit Szenarien in SzenarioDB zu veröffentlichen, wenn dies vom Zuwendungsgeber finanziert wird	42	31
Können sich vorstellen eine Schnittstelle zur SzenarioDB unter bestimmten Voraussetzungen zu entwickeln	58	49
Können sich nicht vorstellen eine Schnittstelle zur SzenarioDB zu entwickeln	5	13
Nutzen regelmäßig Datenbanken wie z.B. Eurostat	64	54
Haben von der OEP gehört und/oder genutzt	87	75

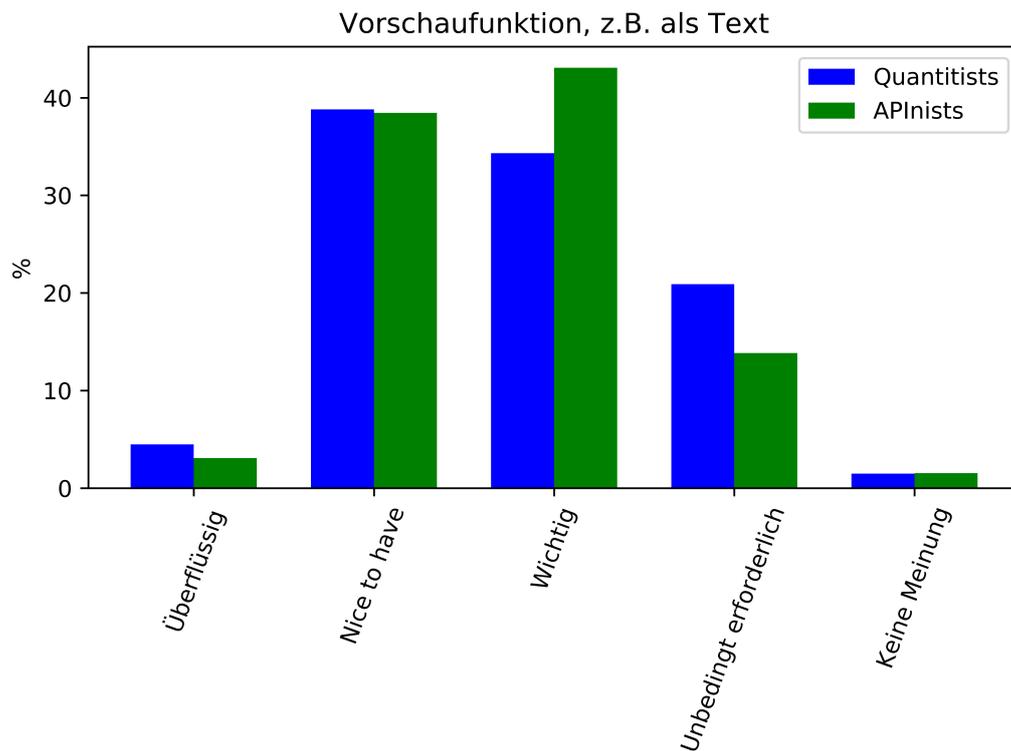


Abbildung 1.6: Einstufung der Wichtigkeit von einer Vorschaufunktion als Tabelle von den beiden Nutzer*innengruppen APInists und Quantitists.

einzig auffällige Abweichung zu der Durchschnittsbefragung (25%) war, dass diese Gruppe zu 46% selten andere Datenbanken verwendet.

Umfrage Stand: 31.08.2018

1.3 Abgleich mit Funktionen der Open Energy Platform (OEP) und Erstellung des Lastenhefts

Neben der Onlineumfrage und den Experteninterviews wurden auch die Nutzer*innen- und Strukturanforderungen aus den bestehenden GitHub-Repositories ausgewertet. Als Stichtag für die GitHub-Issues zur OEP wurde der 29.10.2018 festgelegt. Insgesamt wurden 147 Issues berücksichtigt. Die Issues erfüllen nicht alle unsere oben genannten Anforderungskriterien, da sie in einigen Fällen vor der Entscheidung, User-Stories zu verwenden, und auch von Personen, die nicht am Projekt teilnehmen, erstellt wurden. Insgesamt ergaben sich 27 Anforderungen.

Um Umfage, Interviews und Ausgangs-Issues zusammenzuführen, wurden doppelte Anforderungen entfernt, und jede Anforderung wurde anhand der folgenden Kriteri-

en bewertet: Wichtigkeit, Zeitaufwand, Dringlichkeit und Kategorie. Die Wichtigkeit folgt Schulnoten von eins bis sechs. Die beiden Fragen Zeitaufwand und Dringlichkeit wurden mit der T-Shirt-Größen-Schätzung (9) grob geschätzt (wir haben die Größen wie folgt definiert: S = klein/ein Tag/nicht dringend, M = mittel/eine Woche/wenig dringend und L = groß/ein Monat/sehr dringend). Es waren zwölf Kategorien möglich: IO, Reviewprozess, Bugfixes, Dokumentation, Factsheets, Feature, Funktion zum Ändern am Datensatz, Metadaten, OEP-Layout, Ontologie, Referenzen und Sonstiges. Alle Anforderungen aus der Umfrage wurden ebenfalls in GitHub-Issues übertragen und, wo schlüssig, bestehende Issues zusammengefasst. Das Konzept der Priorisierung und Zeitabschätzung in T-Shirtgrößen wurde übernommen und in Labels an den jeweiligen Issues abgebildet. In der Liste der Issues auf GitHub lassen sich alle zugehörigen Issues mit dem Label `specification_sheet` filtern ³.

Auswertung des Lastenhefts

Aus der Online-Umfrage, den Expertengesprächen und den bestehenden OEP GitHub-Issues wurden insgesamt 69 Anforderungen abgeleitet. Abbildung 1.7 zeigt, dass 43,5% der Anforderungen als sehr dringend und nur 26% als nicht dringend eingestuft wurden. Dies bedeutet, dass ein großer Bedarf an einer Szenariodatenbank und deren spezifischen Anforderungen besteht. Die für die Erfüllung der Anforderungen erforderliche Zeit (grobe Schätzung) beträgt insgesamt 24 Monate. Diese ergibt sich aus 16 Anforderungen mit der Dauer eines Monats, 27 mit der Dauer einer Woche und 25 mit der Dauer eines Tages.

Die Anforderungen liegen am häufigsten (20%) in der Kategorie IO, also Upload und Download von Daten, und in der Kategorie Feature (20%), siehe Abbildung 1.8. Die nächst wichtigsten sind Metadaten (16%), OEP Layout (12%), Funktionen zum Ändern des Datensatzes (9%), Dokumentation (9%) und alle weiteren liegen jeweils unter 5%.

Interessant an den Anforderungen ist, dass sie, obwohl sie sich explizit auf Energieszenarien bezogen haben, nicht nur energiespezifisch sind. Die meisten Anforderungen wären die gleichen für eine Wasserqualitätsdatenbank, oder allgemein formuliert: eine Datenbank zur Speicherung von Modellierungs-Eingangs- und -Ausgangsdaten, wie georeferenzierte Daten, zeitliche Daten und/oder Einzelpunktdateien. Daher kann dieser Ansatz zur Erstellung einer Energieszenariendatenbank auch für andere Disziplinen von Interesse sein.

³https://github.com/OpenEnergyPlatform/oeplatform/issues?q=label%3Aspecification_sheet+

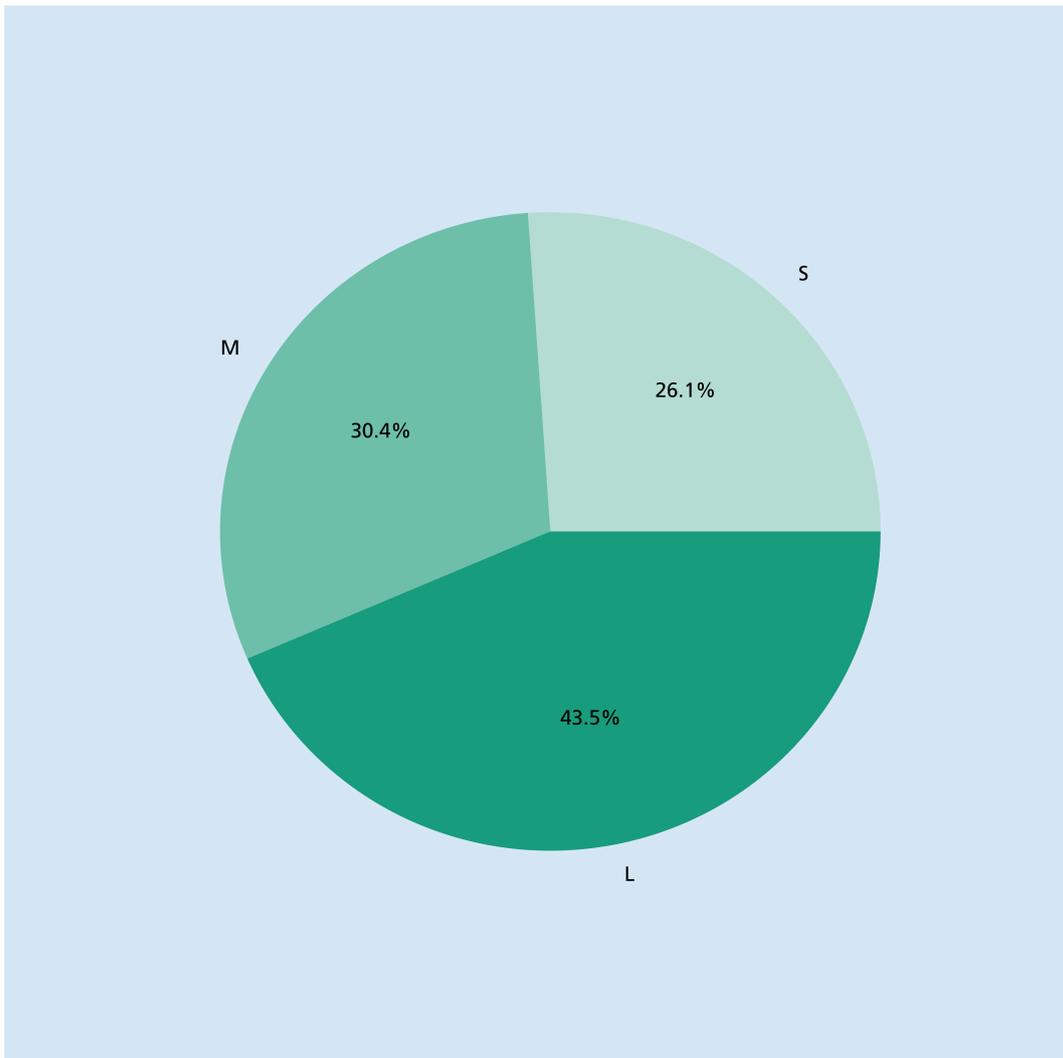


Abbildung 1.7: Der Prozentsatz der Anforderungen, die als sehr dringend (L), mittel dringend (M) und nicht dringend (S) eingestuft wurden.

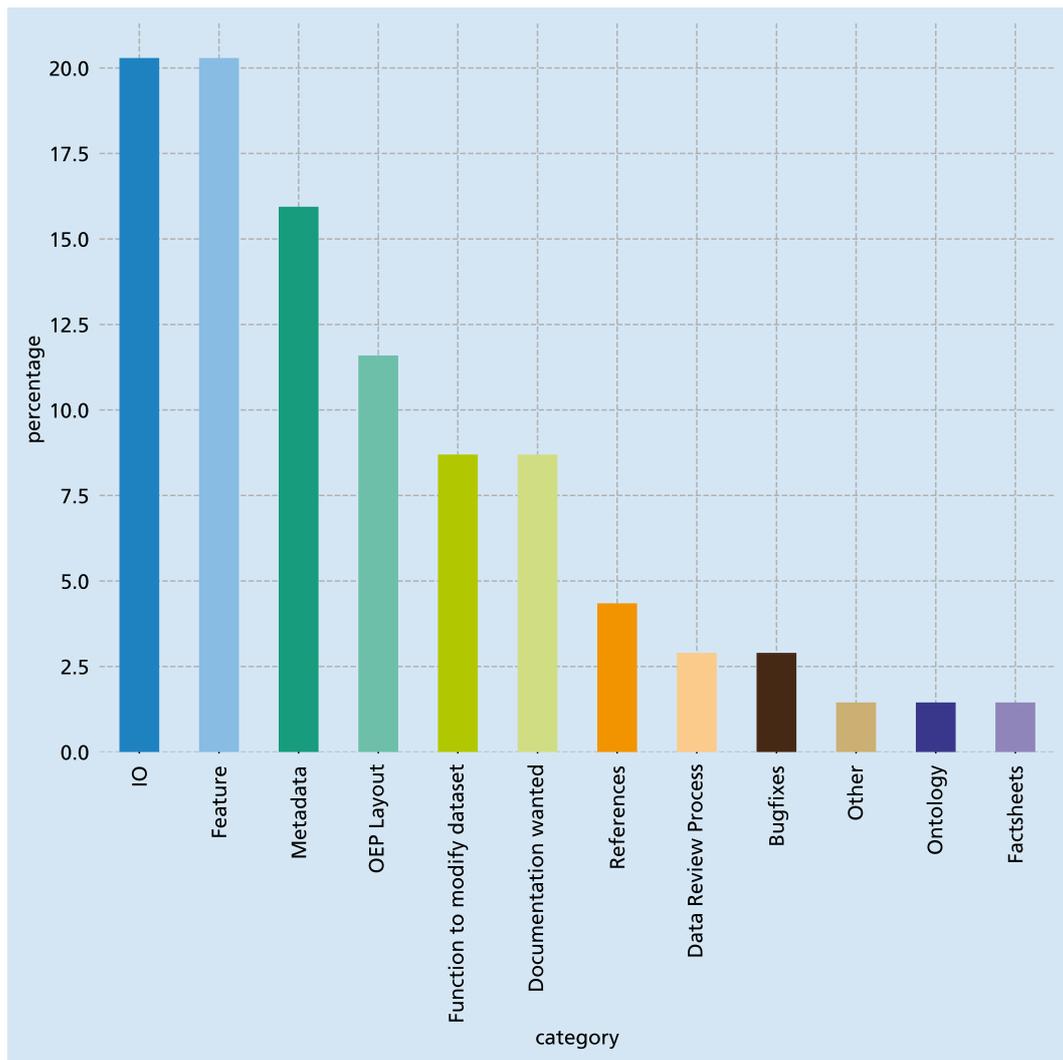


Abbildung 1.8: Anforderungen in Prozent, aufgeteilt in zwölf Kategorien.

2 Ermittlung von Serveranforderungen für zukünftige Nutzung

Die neuen Funktionalitäten der OEP, welche in AP 4 entwickelt wurden, laden Energiesystemmodellierer*Innen ein ihre Energieszenarien mit zu veröffentlichen und so transparenter und zugänglicher zu machen. Zu diesen Szenarien gehören oftmals auch neue Datensätze. Entsprechend mussten angemessene Speicherkapazitäten geschaffen werden.

2.1 Speicherbedarf abschätzen und Datenverkehr hochrechnen

Die Abschätzung des Speicherbedarfs wurde in enger Zusammenarbeit mit dem Forschungsprojekt open_FRED durchgeführt. Im Zuge dieses Projektes wurde die OEP bereits mit großen Mengen an Wetterdaten und Zeitreihen befüllt. Das hierfür erforderliche Speichervolumen belief sich auf 3 TB.

Eine Analyse der Zugriffe auf die OEP, dass der Datenverkehr sehr ungleichmäßig auf verschiedene Datensätze verteilt sind. Einzelne Datensätze wurden verstärkt abgerufen, während andere kaum genutzt werden. Entsprechend dieser Erkenntnisse wurde ein erweitertes Datenbankkonzept entwickelt. Es ist zu erwarten, dass in Zukunft noch weitere große Datenmengen in die Datenbank geladen werden, was zusätzlichen Speicheraufwand bedeutet. Daher wurde der Fokus auf Festplatten gelegt, die eine größere Speicherkapazität, jedoch höhere Zugriffszeiten aufweisen.

2.2 Serverbetrieb

Die bestehenden Server befinden sich in ständigem Betrieb und werden kontinuierlich mit den neusten Sicherheitsupdates versorgt. Im Zuge des Projektes wurden mehrere neue Versionen der Open Energy Platform und der Open Energy Ontology veröffentlicht. Diese und alle notwendigen Pakete wurden auf den entsprechenden

Systemen eingespielt und somit der Öffentlichkeit zugänglich gemacht. Über die gesamte Projektlaufzeit wurden die Log-Dateien der Server überwacht um eventuelle Fehler zu erkennen und zeitnah beheben zu können.

2.3 Hardware-Beschaffung- und Inbetriebnahme

Um das in 4.2 entwickelte Datenbankkonzept umzusetzen, wurde zusätzliche Hardware beschafft. Dem Konzept entsprechend wurde hierbei der Fokus auf Speichermedien mit hoher Kapazität gelegt. Durch die zusätzliche Hardware konnte ein weiterer Server in Betrieb genommen werden, der mit 24 Festplatten ausgestattet ist, von denen jede eine Kapazität von 12 TB aufweist. Teile dieser Kapazität werden verwendet um das Betriebssystem des Servers zu halten und für Redundanzen, die gewährleisten, dass im Falle des Ausfalls einer einzelnen Platte keine Daten verloren gehen. Somit konnte so eine zusätzliche Speicherkapazität von über 100 TB geschaffen werden. Damit ist auch in Zukunft gewährleistet, dass auch historische Daten auf der OEP auffindbar bleiben.

3 Entwicklung des Konzeptes zur Weiterentwicklung der Datenbank

Hauptverantwortung Text: RLI

Auf Basis der durchgeführten Anforderungsanalyse und dem daraus erstellten Lastenheft (siehe 1.3) für eine Szenariendatenbank wurden konkrete Methoden und deren programmiertechnische Umsetzung abgeleitet. Dies teilte sich zunächst in zwei Hauptentwicklungsbereiche: Erstens die Schaffung eines übergeordneten Begriffssystems, die Ontologie, und zweitens die Weiterentwicklung der Open Energy Platform und den dazugehörigen Modulen zum Forschungsdatenmanagement. Dafür wurden bereits existierende technische Ansätze recherchiert, um die gestellten Anforderungen effizient umsetzen zu können. Ein umfassendes Konzept zur praktischen Umsetzung, das Pflichtenheft, ist das wesentliche Ergebnis dieses Arbeitspaketes. Dies beinhaltet auch die Erstellung von UML- und BPMN-Diagrammen, zur Visualisierung der Datenbankstruktur und von Prozessen. Eine wesentliche Grundlage des Open-Science-Ansatzes ist es, Entwicklungen der Forschungscommunity so schnell wie möglich mitzuteilen und Entwicklungen ohne Verzögerungen zur Verfügung zu stellen. Dafür wurden alle Implementationen mit geeigneten offenen Lizenzen auf GitHub gemacht. So ist es möglich transparent zu kommunizieren und Feedback aus der Community direkt aufzunehmen.

Die Konzeptionierung der Programmieraufgaben wurde vom RLI koordiniert und mit allen Projektpartnern kollaborativ entwickelt. Um einen reibungslosen Ablauf der verschiedenen Aufgaben zu gewährleisten, wurden gemeinsame Projektmanagement-Werkzeuge aufgesetzt und gepflegt. In der Praxis kamen zwei Systeme zum Einsatz. Für den Gesamtüberblick auf alle Aufgaben des Projektes wurde ein projektinternes Redmine genutzt. Dieses bildet alle Arbeitspakete als Redmine-Tickets ab und dient primär der Dokumentation der bereits erfolgten Arbeiten im Gesamtkontext mit Zielen, Personenzuweisungen und Fristen. Für die Koordination der Programmieraufgaben wurde die bestehende Organisation der Code-Entwicklung der Open Energy Platform auf GitHub grundlegend überarbeitet und angepasst. Alle Aufgaben wurden als GitHub-Issues angelegt und über die vorhandenen Funktionen strukturiert.

3.1 Ontologie der Energiemodellierungsdomäne

In der Informatik ist eine Ontologie eine formale, (maschinen)lesbare und strukturierte Ablage von Wissen. Das Ziel war es, den Grundstein für den Aufbau einer Ontologie für die Energiesystemmodellierung zu legen und eine erste funktionierende Version zu veröffentlichen. Zu Projektbeginn existierte bereits ein gemeinsames Glossar im Wiki der Openmod-Community⁴. Ein Glossar ist ein Wörterverzeichnis mit einer alphabetischen Liste von Begriffen eines Fachbereichs mit Definitionen und deren Quellen. Ergänzt wurden Abkürzungen und Synonyme der Begriffe. Durch Erweiterungen der Wiki-Software war es zudem möglich, Unter- und Oberbegriffe zu definieren. Zudem ist es möglich zu jedem Begriff einen eigenen Diskussionsfaden zu erstellen. Aus der Community und einer Vorlesungsreihe an der HTW Berlin⁵ wurden 323 Begriffe aus dem Bereich Modellierung von Solaranlagen zusammengetragen. Die so entstandene Sammlung zeigte bereits die Möglichkeiten, aber auch die technischen und strukturellen Einschränkungen auf und diente als Grundlage für die Open Energy Ontology (OEO).

Zu Beginn des Projektes wurden weitere Publikationen identifiziert und gesammelt die wichtige Begriffe für eine Ontologie enthalten. Auf Basis dieser Recherche und Ontologien aus verwandten Wissenschaftsbereichen wurden erste grundlegende strukturelle Entscheidungen zum Aufbau der OEO entwickelt. In mehreren Treffen wurde der theoretische Hintergrund erläutert, diskutiert und die Anpassung an die Energieforschung erarbeitet. Das erste Treffen fand als ganztägiger Workshop an der Uni Magdeburg am 29.05.2018 statt, das zweite am 25.06.2018 als Webkonferenz. Es wurden gemeinsam die Anforderungen und Anwendungsbeispiele an eine Energie-Ontologie gesammelt und aufbereitet. Auf diesen Grundlagen wurde eine erste Testversion der Ontologie erstellt, welche fortan auf GitHub weiterentwickelt wird⁶. Ein drittes Treffen wurde mit allen Projektpartnern am 21.11.2018 durchgeführt. Dabei wurde die grundlegende Ontologiestruktur entwickelt und über die Einbindung existierender Ontologien beraten. Die aus diesen Treffen entstandenen Grundsatzentscheidungen zur Architektur und Struktur der Ontologie werden folgend beschrieben und erläutert.

Die Sprache, in der die OEO geschrieben und formuliert ist, ist die Web Ontology Language (OWL 2) mit der Manchester Syntax⁷. Diese Syntax wurde gewählt, da sie vergleichsweise nutzungsfreundlich ist und ein Überblick von Änderungen in bearbeiteten Dateien schnell möglich. Dies ist für die kollaborative Entwicklung unter der Versionsverwaltung auf GitHub von großem Vorteil. Die OEO baut auf der

⁴<https://wiki.openmod-initiative.org/wiki/Category:Glossary>

⁵<https://re-master.htw-berlin.de/>

⁶<https://github.com/OpenEnergyPlatform/ontology>

⁷<https://www.w3.org/TR/owl2-manchester-syntax/>

Basic Formal Ontology⁸ (BFO) auf. Diese Upper-Level-Ontology definiert grundlegende Konzepte vor und dient als Grundlage für eine Reihe weiterer Ontologien im Wissenschaftsbereich (?). Das führt dazu, dass eine gewisse Kompatibilität zu anderen Ontologien besteht und diese leichter mit der OEO integriert werden können. Es wurde beschlossen, dass die OEO möglichst viele Begriffe übernimmt, die bereits in anderen Ontologien ausgearbeitet wurden. Externe Ontologien, die für die OEO relevant sind, werden fortlaufend im gemeinsamen Wiki des GitHub-Repo dokumentiert.⁹

Die Entwicklung und der Aufbau von Ontologien ist ein langwieriger und andauernder Prozess. Um sowohl eine kontinuierliche Weiterentwicklung zu gewährleisten als auch die Einbettung und Nutzung in der Community zu verankern, wurde auf Erfahrungen aus dem Ontologie-Bereich und der kollaborativen Software-Entwicklung aufgebaut. Es wurden detaillierte Arbeitsabläufe ausgearbeitet und dokumentiert¹⁰. Grundlage dienen die GitHub-Issues, die jeweils einen Diskussionsfaden abbilden (Discussion Workflow). Zur Vereinheitlichung wurden verschiedene Ticket-Vorlagen erstellt, die genutzt werden können¹¹. Die Tickets enthalten damit jeweils eine entsprechende Gliederung und eine Checkliste. Dadurch wird die Bearbeitung effizienter und vereinfacht den Einstieg.

- (A) New ontology term – Hinzufügen eines neuen Begriffes
- (B) Restructure ontology – Umstrukturierung
- (C) Ontology definition update – Verbesserung einer Definition
- Normal issue – Weitere Themen

Neben Absprachen in kleinen Gruppen und kurzen bilateralen Gesprächen ist es nötig länger diskutierte Issues gemeinsam zu besprechen. Zusätzlich zu der schriftlichen Online-Diskussion werden regelmäßige, monatliche Video-Konferenzen angesetzt und durchgeführt. Diese OEO-Developer-Meetings sind ein wichtiger Baustein in der gemeinsamen Ontologie-Entwicklung. Während zunächst die Projektpartner aus SzenarienDB teilgenommen haben, wurden die Treffen um die Partner aus dem LOD-GEOSS-Konsortium erweitert. Die Meetings wurden zunächst vom RLI organisiert und sind grundsätzlich offen für alle Interessierte. Um die Organisation und Vorbereitung besser auf alle Beteiligten zu verteilen, wurde die Planung sorgfältig dokumentiert und eine rotierende Leitung der Treffen eingeführt¹². Die Protokolle der Meetings werden immer gemeinsam in einem EtherPad geführt und anschließend ebenfalls im Wiki archiviert¹³.

⁸<https://basic-formal-ontology.org/>

⁹<https://github.com/OpenEnergyPlatform/ontology/wiki/use-of-external-ontologies>

¹⁰<https://github.com/OpenEnergyPlatform/ontology/wiki/workflow>

¹¹<https://github.com/OpenEnergyPlatform/ontology/issues/new/choose>

¹²<https://github.com/OpenEnergyPlatform/ontology/wiki/OEO-developer-meetings>

¹³<https://github.com/OpenEnergyPlatform/ontology/wiki/OEO-developer-meetings>

Zur technischen Umsetzung der Konzepte kommt die Open-Source-Anwendung Protégé¹⁴ zum Einsatz. Dieses verhältnismäßig komplexe Programm ermöglicht es die Ontologie zu bearbeiten und bieten eine große Bandbreite an zusätzlichen Funktionen. Um den Einstieg für neue NutzerInnen zu erleichtern wurden Anleitungen und Tutorials erstellt, z.B. zum korrekten Einrichten (Setup und Settings) sowie zum Bearbeitungs-Workflow.

Zu Beginn der Ontologie-Entwicklung wurde entschieden ein Steering Committee, das OEO-SC (¹⁵), ins Leben zu rufen. In diesen Lenkungsausschuss wurden renommierte Personen berufen und er setzt sich aus erfahrenen und aktiven ExpertInnen im Bereich der Energiesystemanalyse und der Ontologieentwicklung zusammen. Das Steering Committee soll verschiedene Aufgaben für die OEO wahrnehmen:

- Entwicklung der OEO begleiten
- Förderung der Bekanntheit und Adaption in weiteren Projekten
- Grundlegende Designentscheidungen treffen
- Vernetzung mit anderen Ontologien (z.B. EnArgus)
- Vernetzung in der Energiesystemanalyse-Community
- Schlichtungs- und Entscheidungsinstanz

Das OEO-SC kommt etwa vierteljährlich digital zusammen und wird vom RLI moderiert. Die OEO-EntwicklerInnen berichten über Fortschritte und aktuelle Herausforderungen und das Steering Committee beratschlagt über strategische Fragen. Die Treffen finden unter der Chatham-House-Rule statt, d.h. es werden keine Audioaufnahmen erstellt und Inhalte werden nur ohne Nennung der Identitäten weitergegeben. Das erste Treffen fand am 19.09.2019 statt. Bis Januar 2021 fanden bereits sieben Treffen statt.

Die Veröffentlichung der OEO in Version 1.0.0 fand am 11.06.2020 statt. Dieser Initial Release stellt einen bedeutenden Meilenstein für die Ontologie-Entwicklung dar. Es wurden die bereits beschriebenen grundlegenden Designentscheidungen umgesetzt und die relevanten physikalischen Grundlagen der Ontologie festgelegt. Seitdem wird alle zwei Monate ein neues Release angesetzt¹⁶.

Die OEO wurde in mehreren Workshops der Community vorgestellt. So wurde auf dem Jahrestreffen des Forschungsnetzwerkes Energiesystemanalyse (FNE) 2019 ein erster öffentlicher Workshop zur OEO abgehalten und in weiteren Treffen diskutiert. Weiterhin wurde vom RLI im Rahmen des International FAIR Convergence Symposiums eine Diskussionsrunde zu Metadaten und freier Dateninfrastruktur organisiert, in der die OEO eine zentrale Rolle einnahm.

¹⁴<https://protege.stanford.edu/>

¹⁵<https://openenergy-platform.org/ontology/oeo-steering-committee/>

¹⁶<https://github.com/OpenEnergyPlatform/ontology/releases>

Die Beschreibung und Veröffentlichung der Ontologie in peer-reviewed Journals ist für die Bekanntmachung und Verbreitung in der Wissenschafts-Community von großer Bedeutung. Dabei wird berücksichtigt, dass sowohl den Anforderungen aus der Ontologie-Welt als auch aus dem Forschungsbereich der Energiesystemanalyse gerecht wird. Wie die Ontologie auch, werden die Veröffentlichungen in kollaborativer Zusammenarbeit geschrieben. Zunächst wurde ein Artikel beim *European Knowledge Acquisition Workshop (EKAW)* im Rahmen der *22nd International Conference on Knowledge Engineering and Knowledge Management*¹⁷ eingereicht, der beide Aspekte beinhaltet. Dieses wurde jedoch abgelehnt. Daraufhin wurde entschieden die Inhalte auf zwei spezifische Veröffentlichungen aufzuteilen und es wurden an zwei Publikationen in unterschiedlichen Journals weitergearbeitet. Ein ontologie-technisches Paper wurde bei der *European Semantic Web Conference (ESWC)*¹⁸ eingereicht und fokussiert auf die Beschreibung der Grundstruktur und Evaluierung der entwickelten Konzepte. Das Paper befindet sich aktuell in Review-Prozess. Ein weiteres Paper mit Fokus auf die Energiesystemanalyse-Community wurde im Rahmen einer Sonderreihe für die *Energy Modelling Platform for Europe (EMP-E)*¹⁹ beim Journal *Energy* eingereicht. Dieses wurde jedoch abgelehnt. Eine überarbeitete Version wurde beim Journal *Energy and AI* angestrebt und ist mittlerweile veröffentlicht²⁰.

3.2 Vergleich existierender technischer Ansätze

Für die Weiterentwicklung des Datenbank-Konzeptes wurden bereits existierende technische Ansätze recherchiert und verglichen. Ein Fokus wurde hierbei auf öffentliche Datenplattformen und Datenportale gelegt. Es fanden zudem mehrere Treffen und Gespräche zu technischen Ansätzen mit Betreibern verschiedener Datenbanken statt. Dazu zählen das Joint Research Centre (JRC), die u.a. die *Integrated Database of the European Energy System (JRC-IDEES)*²¹ entwickeln und IIASA die u.a. die Scenario Datenbank des IPCC betreiben²². Zudem wurden weitere projektspezifische Datenbanken aus dem Europäischen Forschungskontext analysiert und ausgewertet. Es sollte geklärt werden, in welchen Projekten bereits an offenen Datenbankinfrastrukturen gearbeitet wird, um Doppelarbeiten zu vermeiden. Eine Zusammenstellung findet sich im Anhang A.3. Leider konnten keine gleichartigen Ansätze einer Community-Datenbank identifiziert werden. Dies und das große Interesse an der Ar-

¹⁷<https://ekaw2020.inf.unibz.it/>

¹⁸<https://www.eswc-conferences.org/>

¹⁹<http://www.energymodellingplatform.eu/>

²⁰<https://doi.org/10.1016/j.egyai.2021.100074>

²¹<https://ec.europa.eu/jrc/en/potencia/jrc-idees>

²²<https://iiasa.ac.at/web/home/research/researchPrograms/Energy/Databases.en.html>

beiten an der OEP zeigten jedoch den Bedarf an den geplanten Entwicklungen der Szenarien-Datenbank.

Das ÖI hat ergänzend die beiden in AP 1.1 am besten bewerteten Datenportale (FAOSTAT und OpenStreetMap) aus technischem Blickwinkel ausgewertet. Das Augenmerk lag dabei auf folgenden Informationen:

- Daten- & Datenbankmodell: Datenbankmodell, Datenbanktechnologie, Datengröße des gesamten "Lagers", Caching, Flexibilität bei Erweiterungen, effiziente Datentypen, Umgang mit unterschiedlichen Datentypen, Implementierung von Metadaten
- Schnittstellen: Import - und Exportschnittstellen
- Benutzung, Verwaltung, QA/QC: Umsetzung der Benutzerverwaltung, Gewährleistung Datenqualität, Möglichkeit der Kollaboration

Die Ergebnisse flossen als Input in den Lastenheftworkshop. Beim 1. Gesamtprojekt-treffen wurden die Kriterien zur technischen Auswertung weiter konkretisiert. Die Auswertung erfolgte dann zum Ende des Jahres 2018. Aus den technischen Analysen und den Betrachtungen zur NutzerInnenfreundlichkeit der bekannten Datenbanken wurde dann ein Datenbank- und NutzerInnen-Konzept für die OEP entwickelt.

3.3 Konzeptionierung

Aufbauend auf den Arbeiten zur Erstellung des Lastenhefts und dem beschriebenen Vergleich existierender technischer Ansätze wurde ein konkretes Konzept zur Weiterentwicklung der Open Energy Platform zu einer Szenarien-Datenbank erarbeitet. Die Kategorisierung der Anforderungen an die Datenbank und das Webinterface nach Dringlichkeit und Wichtigkeit und die Erstellung des Lastenhefts ist in Abschnitt 1.3 ausführlich beschrieben. Ziel war es ein Pflichtenheft zu erstellen, das die aufeinander aufbauenden Arbeiten abbildet, den jeweiligen Partnern zuordnet, eine zeitliche Planung enthält sowie den Fortschritt sichtbar und überprüfbar macht.

Zur Erstellung des Pflichtenhefts wurden alle Issues des Lastenhefts herangezogen und ein Gesamtbild erstellt. Mit der vorläufigen Zeitabschätzung aller geplanten Issues wäre ein Arbeitsaufwand weit über den Rahmen des Projektes hinaus notwendig gewesen. Daher wurden die Aufgaben anhand von Priorität und Abhängigkeiten untereinander ins Pflichtenheft übernommen. Das geschah unter Einbezug einer erneuten Einschätzung aller Konsortialpartner im Rahmen eines Projekttreffens. Das so entstandene Pflichtenheft plante mit der abschätzbaren restlichen Gesamtarbeitszeit bis zum Projektende. Analog zum Lastenheft lassen sich die Issues des Pflichtenhefts

in GitHub nach dem Label “requirement_specification” filtern²³. Das Pflichtenheft selbst liegt zur Archivierung mit allen Kategorisierungen und verlinkten Issues auch in PDF-Form in einem gesonderten Issue²⁴.

Zur Bearbeitung der genannten Anforderungen an das Pflichtenheft wurde die Code-Organisation auf GitHub umgestaltet. So wurden gezielt thematische Repositories angelegt. Etwa wurde ein neues eigenständiges Repository für die Open Energy Ontology angelegt. Der Code für Metadaten, Datenreviews und Tutorials wurde von der Plattform getrennt und ebenfalls in eigene Repositories gelegt. Zur klareren Abgrenzung aller einzelnen Pakete von der OEP, also dem Quelltext für die Webseite und das Webinterface, wurde die Organisation auf GitHub umbenannt. Diese hieß zuvor analog zur Webseite auch Open Energy Platform. Zur besseren Abgrenzung und Kommunikation wurde die *Open Energy Family* als übergeordnetes Framework erschaffen.

Die Organisation innerhalb von GitHub wurde anschließend feiner ausgegliedert. Das Labelling-Schema für Lasten- und Pflichtenheft wurde über alle Repositories hinweg einheitlich gestaltet, Zuweisungen für Programmierarbeiten klar geregelt und einen Austausch direkt im Kontext des Quelltexts ermöglicht. Zudem wurden die Arbeits-, Veröffentlichungs- und Dokumentationsprozesse auf GitHub überarbeitet und dokumentiert. Zunächst wurde ein regelmäßiger monatlicher Termin für neue Releases durchgeführt, um Veränderungen und Weiterentwicklungen schnell öffentlich sichtbar und nutzbar zu machen. Im Projektverlauf wurde der Release-Zyklus dann auf alle zwei Monate festgesetzt, da nun weniger aber größere Neuentwicklungen geplant waren. Zudem wurde eine Testumgebung für Neuentwicklungen an der Open Energy Platform aufgesetzt, um sicher zu stellen, dass neue Entwicklungen auf der laufenden Seite bestehende Funktionen nicht behindern und die Darstellung fehlerfrei ist. TesterInnen wurden dafür als neue Rolle auf GitHub eingerichtet. Die mündlichen Rücksprachen zur Arbeit am Code flossen in die regelmäßigen Besprechungen ein. Bedarfsgleitet wurden Developer-Meetings angesetzt und vorbereitet.

Auf Basis der konzeptionellen Diskussion zu Einheitenumrechnungen und Namenskonventionen erstellten wir Handreichungen und integrierten sie in die OEP:

- Tabelle zur Einheitenumrechnung:
https://openenergy-platform.org/dataedit/view/reference/szenariendb_conversion
- Tabelle zur Einheitenbezeichnung:

²³https://github.com/OpenEnergyPlatform/oeplatform/issues?q=label%3Arequirement_specification

²⁴<https://github.com/OpenEnergyPlatform/oeplatform/issues/394>

https://openenergy-platform.org/dataedit/view/reference/szenariendb_units

- Tabelle mit Global Warming Potentials:

https://openenergy-platform.org/dataedit/view/climate/szenariendb_gwp

Die Ergebnisse dieser konzeptionellen Diskussionen sind dabei in die unter AP 5.3 erstellten Briefings eingeflossen.

Um die heterogenen Strukturen der Landschaft der Energieszenarien und -studien zu erfassen, bedarf es eines flexiblen Datenbanksystems. Um dies zu erreichen wurden mehrere Datenbankkonzepte miteinander verglichen. Diese Analyse ergab, dass ein RDF-Triplestore mit einer Begriffswelt, welche auf den Konzepten und Relationen der Ontologie beruht, am besten geeignet ist um diese Daten zu erfassen. Entsprechend wurde eine Instanz von Apache Jena mit der Fuseki-erweiterung aufgesetzt.

3.4 Softwaretechnische Modellierung im BPMN und UML

Zur Visualisierung von Modellen, Prozessen und Datenstrukturen sind UML- und BPMN-Diagramme ein Standardwerkzeug. Um angelegte Datenmodelle auf der OEP effektiv innerhalb des Konsortiums und nach außen kommunizieren zu können, war es ein Ziel zu identifizieren, welche Darstellungen für welche Bereiche hilfreich sind, wie diese Diagramme automatisiert erstellt werden können und wie diese dargestellt werden.

Das RLI hat für die Erzeugung von UML-Diagrammen verschiedene Programme und Packages analysiert und getestet. Für das Python-Package `iplantuml` ist ein umfangreiches Tutorial erstellt worden²⁵. Dieses wurde innerhalb eines Projekttreffens vorgestellt und auf der OEP zugänglich gemacht. Die Anwendung auf ein bestehendes Datenmodell des oemof-Modells namens `appBBB` ist in Abbildung 3.2 dargestellt.²⁶ Es zeigt für jede Tabelle der Eingangsdaten ein Quadrat, das die wesentlichen Parameter der Datenstruktur enthält: Name der Tabelle, Primärschlüssel, Spaltennamen, Datentypen der Spalten. Zusätzlich sind die Verbindungen und Bezüge zwischen den Tabellen als Striche zwischen den Quadraten dargestellt. Diese werden Fremdschlüssel-Beziehungen genannt.

Zum automatischen Erstellen von UML-Diagrammen aus den Metadaten wurde anschließend das Package `ERAlchemy` getestet und eingebunden. Damit ist es möglich die beschriebenen notwendigen Informationen direkt aus der Datenbank auszulesen

²⁵https://openenergy-platform.org/tutorials/jupyter/UML_Workshop/

²⁶<https://github.com/OpenEnergyPlatform/tutorial/blob/master/omg/appBBB-UML.ipynb>

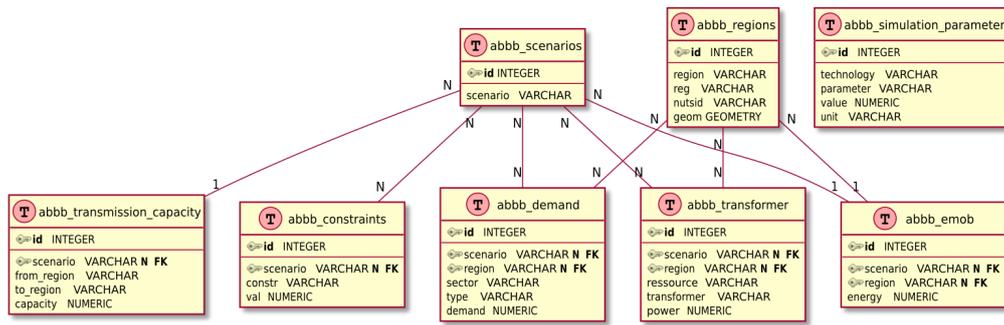


Abbildung 3.1: UML-Diagramm des Datenmodells des Modells appBBB

und darzustellen. Nach einigen technischen Schwierigkeiten konnte eine mit der OEP kompatible Beispielanwendung umgesetzt werden.²⁷

Arbeitsabläufe und Organisations-Prozesse können mit der *Business Process Model and Notation* (BPMN) formalisiert und festgehalten werden. Um mehr Details abbilden zu können, wurde die Konvention erweitert und angepasst. Damit ist es zum Beispiel möglich, den Upload-Prozess von Daten darzustellen und die unterschiedlichen Arbeitsschritte zu zeigen. Der Ablauf des Hochladens eines Datensatzes auf die OEP wird in 3.2 abgebildet. Das Diagramm erfasst die Perspektiven von der hochladenden Person und der Person, die einen Datenreview durchführt. Zusätzlich ist der Ort von Daten und Metadaten zu jedem Zeitpunkt sichtbar.

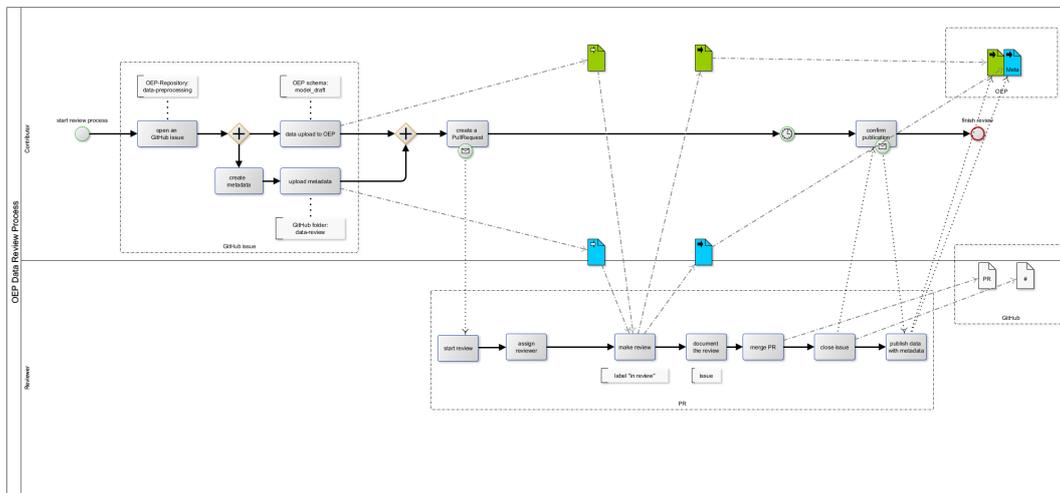


Abbildung 3.2: Business Process Model and Notation (BPMN) für den Review-Prozess beim Hochladens von Daten und Metadaten auf die OEP

Die Darstellung der UML-Diagramme auf der OEP ist vor allem bei der Zusammenstellung von Szenarien-Bundles mit mehreren Daten-Tabellen hilfreich. So kann auf

²⁷<https://github.com/OpenEnergyPlatform/tutorial/blob/develop/omg/UML%20alchemy%20Tutorial.ipynb>

einen Blick gezeigt werden, welche Tabellen auf der OEP zu einem Szenario gehören, wie die Datenstrukturen aussehen und welche Fremdschlüssel bestehen. Die automatischen Visualisierungen werden bei den zukünftigen weitergehenden Umsetzungen in der OEP berücksichtigt.

3.5 Kontinuierlicher Austausch mit der Community

Die KonsortialpartnerInnen waren über den Projektzeitraum hinweg stark in der Energiesystem-Community eingebunden. In der openmod, dem FNE, dem Strommarktverteiler und direkt bei anderen Forschungsprojekten wurde die Open Energy Family aktiv bekannt gemacht, deren Funktionalitäten beschrieben und zur gemeinsamen Entwicklung eingeladen. Insbesondere in der openmod gab es einen regelmäßigen Austausch über den Verteiler, dem Forum und in direkter Korrespondenz. Der gesamte Austausch kann in diesem Rahmen nicht abgebildet werden. Einige öffentlichkeitswirksame Ereignisse sind in Tabelle 3.1 aufgeführt

Zudem wurde aktiv am Forschungsnetzwerk Energiesystemanalyse (FNE) teilgenommen. An den Jahrestreffen hat der Großteil der ProjektmitarbeiterInnen teilgenommen und zahlreiche Inhalte vorgestellt. Über die AG1 Daten und Datenbanken und die AG 5 Transparenz wurde die Open Energy Platform im Netzwerk bekannt gemacht.

Es fand ein kontinuierlicher Austausch auf der etablierten Plattform statt und es konnten die Bekanntheit und der NutzerInnenkreis während der Projektlaufzeit signifikant erweitert werden. Die Open Energy Family hat sich dabei als offene, kollaborative und projektübergreifende Forschungsdateninfrastruktur für die Energiesystemanalyse präsentiert. Sie ist durch das Kommunikations-Konzept sowohl im deutschsprachigen Raum als auch im Europäischen Energieforschungskontext bekannt.

Tabelle 3.1: Öffentlichkeitswirksame Ereignisse in der Energiesystem-Community

Datum	Beschreibung
19.03.2018	Workshop bzw. Expertendiskussion mit dem Netzwerk Suffizienz
21.03.2018	Teilnahme am 3-tägigen, elften Research Data Alliance (RDA) Plenary: https://www.rd-alliance.org/plenaries/rda-eleventh-plenary-meeting-berlin-germany/rda-11th-plenary-programme
19.04.2018	Vorstellung von Postern zur OEP auf dem Jahrestreffen des Forschungsnetzwerks Energiesystemanalyse in Berlin https://reiner-lemoine-institut.de/poster-the-openenergy-platform-oep-und-the-openenergy-database-oedb
07.06.2018	Vorstellung von Postern zur OEP und OEDB auf der openmod in Zürich https://reiner-lemoine-institut.de/poster-the-openenergy-platform-oep-und-the-openenergy-database-oedb
30.06.2018	Berlin: BMWi-Workshop zu Open Data for Electricity Modelling (Öko, RLI)
12.09.2018	Berlin: ESYS Energiesymposium: ENERGI3: Integriert, intelligent, international – die nächste Phase der Energiewende (Öko)
24.09.2018	Karlsruhe: Energy Scenarios Conference (ESS Conference 2018) (Öko, IEE, OvGU)
16.-17.10.2018	Berlin: FVEE Jahrestagung (Öko)
30.10.2018	Berlin: Abschlussworkshop des open-eGo-Projektes (Öko, OvGU, RLI)
19.02.2019- 20.02.2019	Potsdam: RDA Deutschland Tagung 2019 (Öko, RLI)
22.-24.05.2019	Aarhus: OpenMod Konferenz (Öko, OvGU)
17.06.2019	Berlin: NGO-Open Modeling Community "Bridge Meeting"(Öko)
16-24.09.2019	Bozen-Bolzano: Bolzano Rules and Artificial Intelligence Summit (OvGU)
19.09.2019	Kick-Off des OEO Steering Committees als Repräsentanz der Community und moderiert von den Konsortialpartnern. Nachfolgend vierteljährliche Termine (RLI)

Tabelle 3.2: Öffentlichkeitswirksame Ereignisse in der Energiesystem-Community Part II

Datum	Beschreibung
26.09.2019	Vorstellung des OEP-Metadaten-Standards im FNE https://www.youtube.com/watch?v=hpXx6zp5gZ0
21.10.2019	Helsinki: Annual semantic interoperability conference (SEMIC)(RLI)
23.10.2019	Vortrag auf der 21. International Conference on Grey Literature (GL21) in Hannover https://reiner-lemoine-institut.de/open-science-collaborative-working-energy-system-research-huelk-2019/
08.11.2019	Jülich: Ergebnisworkshop zum Rechtsgutachten für Monitoringdaten (RLI)
15.11.2019	STADT: Abschlussworkshop des DemandRegio Projektes, (RLI)
01.04.2020	Schreiben zur OEP in der openmod (https://groups.google.com/g/openmod-initiative/c/XaN2VC1_0_s)
13.05.2020	Lightning Talk in der openmod zur OEO (https://www.youtube.com/watch?v=Cfz8Bdu0mic)
14.05.2020	Workshop zur OEP auf dem oemof-Developer-Meeting: https://github.com/oemof/oemof/wiki/Meeting-May-2020-%28online%29
06.10.2020	EMP-E Project Session SzenarienDB (IEE, Öko, RLI). Einleitungsvideo: https://www.youtube.com/watch?v=rp9ff-ZekLE
07.10.2020	EMP-E Poster Session zur OEO. Einleitungsvideo: https://www.youtube.com/watch?v=TRJTXZNondo
02.12.2020	Diskussion auf dem International FAIR convergence Symposium zu offener Infrastruktur und Ontologie

4 Programmiertechnische Umsetzung des Konzeptes

Hauptverantwortung Text: OvGU Allgemeines zu diesem AP

4.1 Erweiterung des Open-Source-Konzeptes

Bei Entwicklung der Open Energy Family ist die Offenheit und Transparenz ein wichtiges Kriterium. Dies bedeutet nicht nur, dass Plattformen und Wege geschaffen werden, mit denen EnergiesystemmodelliererInnen ihre Daten, Modelle und Prozesse transparent dokumentieren und sichtbar machen können, sondern auch, dass die Entwicklungsprozesse der Open Energy Family selbst diesen Werten folgen. Dementsprechend wird die Entwicklung auch weiterhin in einer öffentlich zugänglichen Github-Organisation durchgeführt. Dies ermöglicht es Interessierten zu verfolgen, welche nützlichen Funktionalitäten derzeit in Entwicklung sind und wann diese veröffentlicht voraussichtlich werden. Darüber hinaus können die Entwicklungsprozesse auch durch das Anlegen neuer Issues jederzeit beeinflusst werden. Somit ist ein ständiger Austausch mit den EndanwenderInnen und anderen Interessierten bezüglich ihrer Erwartungen und Bedürfnisse. Um diese Austauschprozesse einfacher zu gestalten, wurden die Arbeitsabläufe im Zuge des Projektes neu gestaltet und genau dokumentiert. Im folgenden werden diese Neuerungen anhand eines beispielhaften Anwendungsfalls dargelegt.

Eine NutzerIn findet einen Fehler in einem der Module der Open Energy Family - z.B. der Open Energy Platform. Über entsprechende Weiterleitungen auf der Open Energy Platform wird die Betreffende auf die entsprechenden Seiten in den Github-Repositories weitergeleitet. Hier wurden Labels und wie etwa im Ontologie Repository Vorlagen für Issues angelegt²⁸, die die Beschreibung bestehender Probleme erleichtern und es den Entwicklern erleichtert entstandene Probleme nachvollziehen zu können. Nachdem das Ticket erfolgreich erstellt wurde, wird es in den Entwicklungszyklus aufgenommen und mit einer voraussichtlichen Zielversion versehen und dem entsprechenden Github-Versionsprojekt zugewiesen. Hierdurch wird

²⁸<https://github.com/OpenEnergyPlatform/ontology/issues/new/choose>

sichergestellt, dass jederzeit klar dokumentiert ist, welche Tickets sich momentan in Bearbeitung befinden, bzw. welche noch zu bearbeiten sind oder bereits bearbeitet wurden. Nachdem das beschriebene Problem behoben oder das gewünschte Feature implementiert wurde, wird es in den nächsten geplanten Release aufgenommen. Für die Open Energy Platform wurde ein zweimonatiger Release-Zyklus vereinbart - d.h. in jedem zweiten Monat wird eine neue Version der Plattform veröffentlicht und auf den Live-Servern eingespielt. Zu jeder neuen Version wird auch eine lesbare Zusammenfassung der wichtigsten Änderungen erstellt, die es Nutzern ermöglicht sich schnell eine Übersicht über neue Funktionalitäten zu verschaffen.

Somit wurde sichergestellt, dass die bestehenden Arbeitsabläufe für nachvollziehbar und transparent sind.

4.2 Entwicklung Datenbankstruktur

Die Analyse der bestehenden Datensätze hat gezeigt, dass den Datensätze in der Energiesystemmodellierung überaus heterogene Strukturen zugrunde liegen. Wie sich im Laufe der Projektbearbeitung und Stakeholderkommunikation herausstellte, sind diese Strukturen auch notwendig, da diese häufig auf bestimmte Energiesystemmodelle zugeschnitten sind und die Transformation in eine einheitlich festgelegtes Datenformat oft überaus rechen- und ressourcenaufwendig ist. Durch das Erzwingen solch fester Strukturen würde die Nutzung vieler bestehender Energiesystemmodelle erschwert, was die Nutzbarkeit und Akzeptanz sowohl der Modelle als auch der Open Energy Family behindern würde. Daher wurde ein im Zuge des Projekts auf einen flexibleren Ansatz gesetzt, der es erlaubt nutzerdefinierte relationale Datenmodelle zu verwenden und diese mit entsprechenden Metadaten zu annotieren.

Dennoch wurde ein empfohlenes Datenformat skizziert und entworfen, dass eine besonders übersichtliche und effiziente Ablage der Daten erlaubt. Das OEDatamodel wurde in Zusammenarbeit der Forschungsprojekte `open_MODEX` und `SzenarienDB` entwickelt und implementiert.²⁹ Es bietet eine Möglichkeit, Daten in einem vorgefertigten Format als `Datapackage` und Datenbank-Tabellen abzulegen. Dies erweitert die absichtlich komplett flexibel gehaltene Datenstruktur der OEP um einen potentiellen Standard.

Die Beschreibungen der Szenarien sind ebenso heterogen wie die genutzten Daten. Um die Szenarien dennoch in ihrer Komplexität abbilden zu können wurde parallel zur relationalen Datenbank der OEP ein RDF-Triplestore eingerichtet. Diese Technologie wurde bereits erfolgreich zum Erfassen komplexer heterogener Daten genutzt

²⁹<https://github.com/OpenEnergyPlatform/oedatamodel>

(1). Begriffe und Relationen der in 3.1 erstellten Ontologie dienten als Grundlage für die Struktur dieser Datenbank. Dieses Vorgehen legt den Grundstein für eine Zusammenführung verschiedenster Konzepte der Energiesystemmodellierung. So können Szenarien in Zusammenhang mit den verbundenen Ein- und Ausgangsdaten sowie involvierten Instituten, Fördermittelgebern und wissenschaftlichen Veröffentlichungen betrachtet werden.

Zuletzt wurden auch eine neue Art von Factsheet für Energieszenarien und -studien erstellt. Diese basieren auf einer Wissensdatenbank, welche auf der in 3.1 erstellten Ontologie beruhen. Hierbei werden Daten in Dreierpaaren - sogenannten Triples - abgespeichert. So kann der Fakt, dass SIROP vom BMWI gefördert wurde wie folgt ausgedrückt werden:

```
:SIROP OEO:00000509 :BMWI .
```

Der numerische Identifikator steht hierbei für die Relation "has funding source", welche in der OEO definiert wurde. Diese Struktur wurde bereits in der Vergangenheit z.B. durch DBPedia (1) genutzt um komplexe Zusammenhänge flexibel erfassen zu können. Anhand der Szenarienraster, die im Projekt erstellt wurden, wurden geeignete Relationen und Konzepte aus der Ontologie ausgewählt um einen Wissensgraphen für Szenarien und Studien zu definieren.

4.3 Entwicklung und Optimierung der Datenbankschnittstellen

Die Analysen der Datenbank- und Anfragezeiten haben gezeigt, dass die größten Flaschenhälse im Serialisieren der Daten Zwischen der Datenbank und dem Nutzer lagen. Um diese zu beheben, wurden die Schnittstellen der Plattform um moderne Streaming-Mechanismen erweitert. Diese erlauben die Datenverarbeitung, während noch weitere Daten empfangen werden.

Eine einzelne Datenbank kann nur eine begrenzte Anzahl an Anfragen verarbeiten – insbesondere wenn größere Datenmengen abgefragt werden. Um die steigenden Nutzeranfragen dennoch reibungslos verarbeiten zu können, wurde die Sitzungsverwaltung der OEP überarbeitet. So ist es nun möglich, dass jede NutzerIn seinen eigenen Pool an Verbindungen verwalten kann. Hierdurch wird sichergestellt, dass eine einzelne Nutzerin nicht den durch viele parallelisierte Abfragen das Nutzungserlebnis anderer Nutzer beeinträchtigen.

Eine Analyse des Abfrageverhaltens im Zuge von AP ?? hat gezeigt, dass sich die Anfragen nicht gleichmäßig auf die Datensätze in der Datenbank verteilen. Stattdessen wurden vermehrt Anfragen an die Daten aus den Projekten open_eGo. Dies ist insbesondere dadurch zu erklären, dass die dort entwickelten Tools in enger Zusammenarbeit mit der OEP stehen. So greifen Werkzeuge, wie *eTraGo* direkt auf die Daten der Datenbank zu. Auch Daten aus dem Projekt openFRED wurden häufig abgerufen. Entsprechend dieser Beobachtung wurde ein erweitertes Datenbankkonzept entwickelt und im Zuge dieses APs umgesetzt. Hierbei wurde die Datenbank um einen sogenannten *Cold-Storage* erweitert. Hierbei handelt es sich um ein zweischichtiges Datenbanksystem. Dieses besteht aus einem schnellen Speicher mit limitierter Kapazität und einem langsamerem Speicher mit großer Speicherkapazität. Datensätze, die weniger häufig abgefragt werden, werden in den langsameren Speicher verschoben, um neuen Platz für hochfrequentierte Daten zu schaffen. Somit wird gewährleistet, dass die Datenbank gute amortisierte Abfragezeiten aufweisen kann.

4.4 Erweiterung der Open Energy Platform (OEP)

Wie bereits in der Konzeptionierung beschrieben wurde bei der Erweiterung der Plattform nicht nur das Webinterface und die dahinter liegende Datenbank betrachtet, sondern auch daran gearbeitet die Open Energy Family als modulares und offenes Framework für kollaboratives Datenmanagement auszubauen. Aus den gesammelten Anforderungen und den praktischen Erfahrungen mit der Nutzung der bestehenden Tools und Pakete wurde ersichtlich, dass an verschiedenen Modulen parallel entwickelt werden sollte. So entstand ein aufeinander abgestimmter Baukasten. Während einige Familienmitglieder bereits bekannt und etabliert sind, gibt es andere die eher unscheinbar sind. Die meisten haben ein eigenes Repository (z.B. die OEMetadata, OEO), manche sind direkter Bestandteil eines anderen Repositories (z.B. Factsheet, OEDB) und einige werden nicht explizit genannt und stellen zusätzliche Funktionalitäten bereit (z.B. oem2orm, open-MaStR).

Für eine schnelle Anwendung von Funktionalitäten sind gut strukturierte und einfach auffindbare Tutorials zu den verschiedenen Modulen der Open Energy Family von besonderer Bedeutung. Die bereits bestehenden Tutorials lagen überwiegend als Jupyter Notebooks mit Beispiel-Code im GitHub. Um diese mit weiteren Formaten zu ergänzen, wurde eine neue Tutorial-Section auf der OEP geschaffen.³⁰ Dieses bietet zum einen die Möglichkeit, die vorhandenen Notebooks in die Webseite einzubinden als auch neue Tutorial in einem einfachen Editor im Markdown-Stil

³⁰<https://openenergy-platform.org/tutorials/>

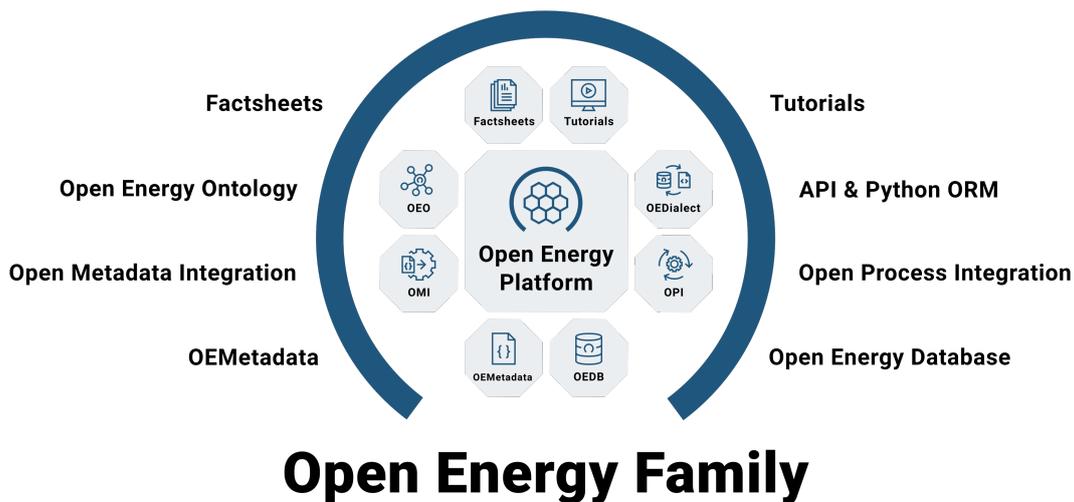


Abbildung 4.1: Darstellung der Module der Open Energy Family

zu erstellen. Dadurch können auch andere Formate, wie Texte, Bilder und sogar Videos eingebunden und dargestellt werden. Eine zusätzliche Kategorisierung und Sortierung der Tutorials ermöglicht eine wesentlich bessere Übersicht und erleichtert das Auffinden der passenden Seiten.

In Kooperation mit der Openmod-Community und nach verschiedenen Workshops wurden die OEMetadaten überarbeitet und versioniert veröffentlicht.³¹ Dabei konnte auf die praktische Umsetzung in verschiedenen Forschungsprojekten an unterschiedlichen Datensätzen aus der Energiesystemmodellierung (Kraftwerkslisten, Wetterdaten, Parametersammlungen, Kennlinien, Einspeisezeitreihen und viele weitere) zurückgegriffen werden. Auch konnten verschiedene Anforderungen aus Auftragsprojekten eingearbeitet werden, z.B. die Nennung des Projektkontextes und die Darstellung des Publisher-Logos. Zudem wurde die Dokumentation und die Beschreibung der Felder wesentlich verbessert.³²

Zum einfacheren Umschreiben und Validieren der Metadaten wurde das Python-Paket Open Metadata Integration (OMI) entwickelt und veröffentlicht. Dies ermöglicht es auf der OEP und als eigenständiges Tool die Metadaten zwischen den verschiedenen Versionen so weit wie möglich automatisch umzuwandeln sowie umfangreiche Tests zur Validierung durchzuführen.³³

Um der OEP ein neues, ansprechendes Aussehen (vgl. Abb. 4.2) zu verleihen, wurden die bestehenden Icons und Logos grundlegend überarbeitet und ergänzt. In Zusammenarbeit mit einem Grafikdesigner, der kurzzeitig das Projektteam am RLI

³¹<https://github.com/OpenEnergyPlatform/oemetadata/releases>

³²https://github.com/OpenEnergyPlatform/oemetadata/blob/develop/metadata/v141/metadata_key_description.md

³³<https://github.com/OpenEnergyPlatform/omi>

unterstützt hat, wurde auch ein neues Grunddesign für die Open Energy Family entworfen. Dieses hebt sich durch einen deutlich dünneren Kreis von den bestehenden Projektlogos und dem alten OEP-Logo ab. Die Logos und Icons wurden sowohl direkt auf der OEP, z.B. der Startseite, verwendet als auch in den GitHub-Repositories.³⁴ Der Aufbau und die Struktur der Webseite wurden ebenfalls überarbeitet (vgl. Abb. 4.3 und Abb. 4.4)

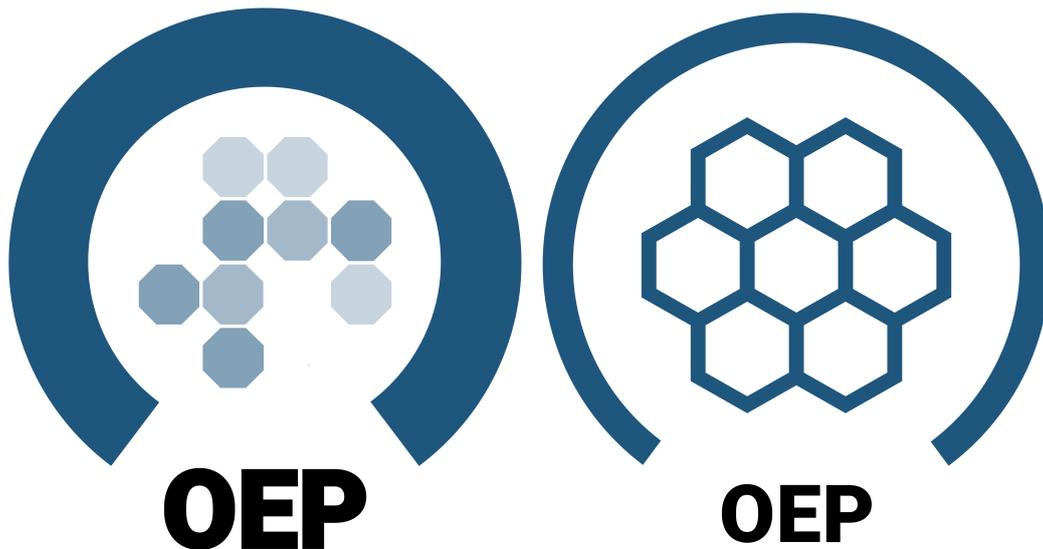


Abbildung 4.2: Das alte und das neue Logo der OEP (von links nach rechts)

Um den Upload- und Download-Prozess über den bestehenden OEDialect zu vereinfachen, wurden zwei Python-Pakete entwickelt, das aus bestehenden Metadaten oder Datenbanktabellen automatisch die objektrelationale Abbildung, die object-relational mapping (ORM) extrahieren. Diese Objekt-Klassen können dienen als standardisierte und funktionale Schnittstellen zwischen den Modulen und der Datenbank und sind für erfahrene ProgrammiererInnen von großer Wichtigkeit. Das erste Paket heisst *saio*³⁵ und leitet sich von SQLAlchemyInputOutput ab. Es ist eine Erweiterung der automatischen Tabellendefinition von SQLAlchemy und erlaubt das generieren eines ORM aus einer bestehenden Tabelle in der OEP. Einfach gesagt wird nur mit Angabe des Schemas und des Tabellennames die Tabellenstruktur übergeben. Als weiterer Baustein dient das Paket *oem2orm*³⁶ das aus den ausgefüllten Metadaten, genauer gesagt den dokumentierten Ressourcen-Beschreibungen, den ORM generiert. Die OEMetadaten enthalten nämlich ebenfalls alle relevanten Information. Dadurch kann aus den korrekt ausgefüllten Metadaten (z.B. für eine CSV-Datei) die Tabelle auf der OEP angelegt werden in die die Daten dann geladen werden können.

³⁴<https://github.com/OpenEnergyPlatform/organisation>

³⁵<https://github.com/coroa/saio>

³⁶<https://github.com/OpenEnergyPlatform/oem2orm>

4.4 Erweiterung der Open Energy Platform (OEP)

Abbildung 4.3: Design der OEP vor dem Redesign

Abbildung 4.4: Design der OEP nach dem Redesign

Im Rahmen des Projektes wurde eine Masterarbeit angefertigt, die die durchgeführten Weiterentwicklungen der Szenarien-Datenbank in Bezug auf die identifizierten NutzerInnenanforderungen evaluiert. Es wurde der Status Quo begutachtet und dokumentiert, mit den identifizierten Anforderungen abgeglichen und daraus weiterführende Verbesserungen abgeleitet.(10)

Es wurden verschiedene Usability-Tests an der OEP durchgeführt, um Probleme und Resultate der im Projekt entstandenen Forschungsarbeit zu beurteilen. Zu diesem Zweck wurde eine angemessene und vergleichbare Evaluationsmethode ausgewählt, die sowohl aus einer empirischen als auch einer analytischen Komponente besteht. Im empirischen Teil wurde ein Usability-Test unter Zuhilfenahme von Observation, Videoaufzeichnung, Technik des lauten Denkens und ein Fragebogen mit insgesamt 14 Probanden durchgeführt. Im analytischen Teil wurde eine Evaluierung mit der Inspektionsmethode „Streamline Cognitive Walkthrough“ mit 2 fiktiven Nutzungsprofilen vollzogen.

Die im Rahmen der Masterarbeit durchgeführten Untersuchungen zeigten, dass es in Bezug auf die geforderten Anforderungen noch einige Schwächen bei der Nutzbarkeit der Szenariendatenbank gab. Daraus wurde ein Empfehlungsbogen zusammengestellt, der die gefundenen Probleme auflistet. Die gruppierten und eingestuften Handlungsempfehlungen werden in Abbildung 4.5 dargestellt. Es zeigt einen Fokus auf der Nutzungsoberfläche sowie der Informationsarchitektur der Inhalte auf der OEP. Als Konsequenz fallen Verbesserungsmöglichkeiten in den Bereichen Verknüpfungen und Verwaltungsmanagement von Datensätzen auf.

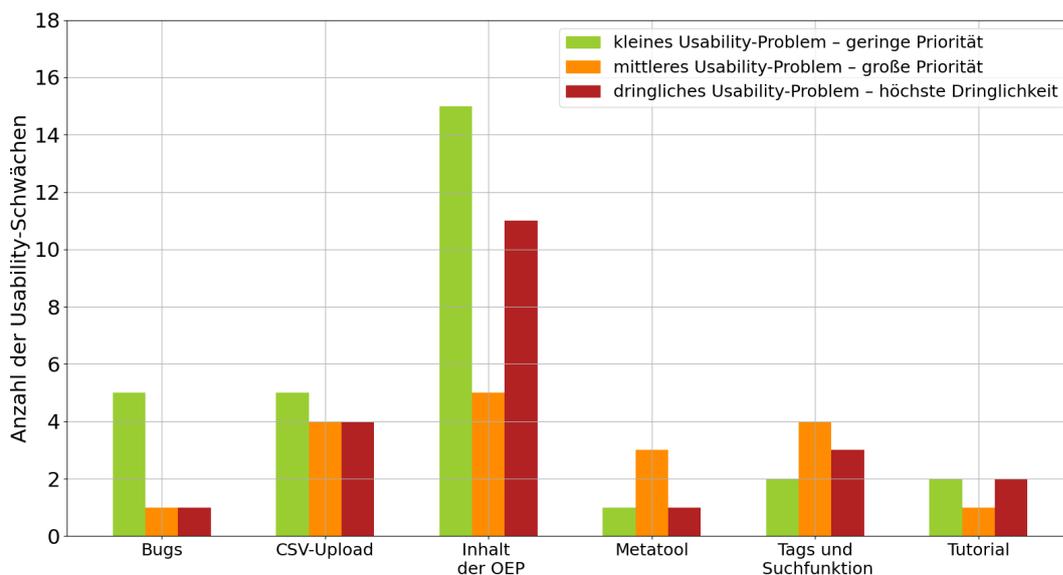


Abbildung 4.5: Zusammenfassung aller erfassten Usability-Schwächen (10, S.62)

Aus den Ergebnissen der Untersuchungen wird zudem deutlich, dass die Konzeptionierung der OEP als eine offene, zugängliche Datenbank wahrgenommen wird. Demnach spricht die Bedienoberfläche primär Zielgruppen an, die Akteure im Interessensfeld der Ingenieursdisziplin, Programmierung und Forschung vertreten. Darüber hinaus weist es überwiegend eine intuitive Struktur für Fachfremde auf. Die größte Erkenntnis des Usability-Test ist, dass die Qualität und der inhaltliche Wert der Datenbestände die höchste Relevanz hat. Dieser Aspekt begründet maßgeblich die Wiederkehrende Benutzung der Plattform. Es wird in Zukunft intensiv daran gearbeitet, gut dokumentierte und kuratierte Datensätze auf der OEP bereitzustellen.

From the changelogs:

- Show changelogs
- Support alembic migrations
- Add contribute file (#469, #505)
- Subpage for OEO Steering committee
- Add 'form=csv' option for /rows-API-requests
- API: More fine grained handling of temporal data
- API: Explicit column handling (#484)
- API: Handle metadata via API
- API: Pass GIS-types in column descriptions (#507)
- Metadata upload and download via API (#461)
- Add tutorials app (#489)
- Show video-content in tutorials (#516)
- Factsheets: General remodeling of factsheets (#419)
- Factsheets: Make links clickable (#74)
- Add button for CSV download
- Support metadata version 1.4
- Allow Metadata on table creation

Der oben beschriebene Wissensgraph diene als Grundlage für eine neue Art von Factsheets auf der OEP. Diese bieten Formulare, welche die verschiedenen Relationen aus der Ontologie repräsentieren. So werden definierte Energieträger, und Sektoren direkt aus der Ontologie geladen und als Optionen in den entsprechenden Feldern zur Verfügung gestellt.

5 Szenariendatenrecherche und exemplarische Befüllung der Datenbank

Eine exemplarische Befüllung der OEP mit offen lizenzierten Szenariendaten war das primäre Ziel dieses Arbeitspaketes. Die ausgewählten Beispiel-Szenarien dienen der praktischen Umsetzung der technischen und juristischen Arbeitsschritte zur Veröffentlichung und Dokumentation.

Zunächst schaffte die Szenarienrecherche einen Überblick darüber, welche Daten aus einer Auswahl von veröffentlichten Szenarien prinzipiell vorlagen und unter welchen Bedingungen diese auf der OEP der Öffentlichkeit verfügbar gemacht werden können (AP 5.1). Danach widmete sich das AP 5.2 der Prüfung der Verfügbarkeit der Daten für die Datenbank. Implikationen für Datenformate war das Thema des AP 5.3. An dieser Stelle erstellten wir zwei Briefings, kombiniert mit Handreichungen in der OEP (siehe AP 3.3). In AP 5.4 kümmerten wir uns um Fragen der Datenlizenzierung, die während der Arbeiten anfielen. Nach der Klärung der Datenlizenzierungen integrierten wir in AP 5.5. die Szenariendaten – mit offenen Lizenzen versehen – in die OEP.

Im Folgenden dokumentieren wir den Verlauf und die Ergebnisse detailliert je Unterarbeitspaket.

5.1 Recherche Szenarien

Wir sichteten eine Auswahl an veröffentlichten Szenarien und analysierten die in ihnen dokumentierten Datensätze. Für diesen Zweck erarbeiteten wir ein Erhebungsraster, in dem wir Szenarien eintrugen und spezifizierten. Je Studie wurden Informationen zu folgenden Punkten erhoben (das detaillierte Raster befindet sich im Annex A.4) beigefügt.

- General information about study
- Publication format and license
- Analysis space
- Resolution
- Modelling results
- Exogenous data (modelling input)
- Endogenous data (calculated by model)

Je Szenario aus jeder Studie erhoben wir folgende Informationen:

- Scenario name
- Scenario abbreviaton
- Definition and target
- Target-driven scenario?
- Measure-driven scenario?
- New aspects
- Type of sensitivities

Insgesamt bestückten wir das Raster mit folgenden Szenarien:

1. Aus der Studie *Klimaschutzszenario 2050* (2. Runde) mit:
 - Aktuelle-Maßnahmen-Szenario
 - Klimaschutzszenario 80
 - Klimaschutzszenario 95
2. Aus der Studie *Klimapfade für Deutschland* mit:
 - Referenzszenario
 - Globaler Klimaschutz – 80% Klimapfad
 - Globaler Klimaschutz – 90 % Klimapfad
 - Nationale Alleingänge – 80% Klimapfad
 - Nationale Alleingänge – 95% Klimapfad
3. Aus der Studie *Zukunft Stromsystem* mit:
 - Referenz-Szenario
 - Schnellausstiegs-Szenario
 - Laufzeitbegrenzungs-Szenario – 20 Jahre (EEG)
 - Laufzeitbegrenzungs-Szenario – 25 Jahre (EEG)
 - Laufzeitbegrenzungs-Szenario – 30 Jahre (EEG)
 - Laufzeitbegrenzungs-Szenario – 20 Jahre
 - Laufzeitbegrenzungs-Szenario – 25 Jahre
 - Laufzeitbegrenzungs-Szenario – 30 Jahre
 - Transformations-Szenario
4. Aus der Studie *Verbesserung der methodischen Grundlagen und Erstellung eines Treibhausgasemissionsszenarios als Grundlage für den Projektionsbericht 2017 im Rahmen des EU Treibhausgasmonitorings („Politikszenerien VIII“)* mit:
 - Mit-Maßnahmen-Szenario
 - Mit-weiteren-Maßnahmen-Szenario
5. Aus der Studie *Untersuchungen zur Energiestrategie Brandenburgs* mit:
 - Energieszenario 2030
 - Grüne 2030
6. Aus dem Modell *oemof-B3 - Analysis of the energy system of Brandenburg and Berlin*:
 - Szenario 1
 - Szenario 2
 - Szenario 3

7. Aus der Studie *Transformationsprozess zum treibhausgasneutralen und ressourcenschonenden Deutschland* mit:

- Energy efficiency
- Late development
- Material efficiency
- Lifestyle
- Supreme

8. Aus der Studie *Transformationspfade im Wärmesektor: Betriebs- und volkswirtschaftliche Betrachtung der Sektorkopplung mit dem Fokus Fernwärme mit hohen Anteilen konventioneller KWK Erzeugung und Rückkopplung zum Gesamtenergieversorgungssystem*³⁷ mit insgesamt 4 Szenarien:

- Szenario mit Einsatz von viel Biomasse und einem hohen Sanierungsgrad im Gebäudesektor in Deutschland
- Szenario mit Einsatz von viel Biomasse und einem moderaten Sanierungsgrad im Gebäudesektor in Deutschland
- Szenario mit Einsatz von wenig Biomasse und einem hohen Sanierungsgrad im Gebäudesektor in Deutschland
- Szenario mit Einsatz von wenig Biomasse und einem moderaten Sanierungsgrad im Gebäudesektor in Deutschland

Szenarienauswahl In die OEP integriert werden sollten mindestens folgende der oben gelisteten Szenarien: Das Klimaschutzszenario 95 (KS 95), beide Szenarien aus Untersuchungen zur Energiestrategie Brandenburg, und Daten aus dem BMWi-geförderten Projekt *Transformationspfade im Wärmesektor: Betriebs- und volkswirtschaftliche Betrachtung der Sektorkopplung mit dem Fokus Fernwärme mit hohen Anteilen konventioneller KWK-Erzeugung und Rückkopplung zum Gesamtenergieversorgungssystem*

Dokumentation Datensätze Eine hinreichend strukturierte Dokumentation der verfügbaren Datensätze erfolgte durch Hochladen der Daten in die OEP und gleichzeitiger Dokumentation durch den im Projekt weiter entwickelten Metadatenstandard. Alle zum Upload in die Datenbank identifizierten Szenarien lagen schon zu Prozessbeginn in maschinenlesbarem Format vor und mussten daher nicht separat nach Art dokumentiert werden. Durch Abfrage der in die Datenbank integrierten Szenariendaten kann daher direkt über die OEP ein strukturierter Überblick der vorhandenen Datensätze generiert werden.

³⁷https://www.iee.fraunhofer.de/content/dam/iee/energiesystemtechnik/de/Dokumente/Veroeffentlichungen/2019/2019_Feb_Bericht_Fraunhofer_IEE_-_Transformation_Waerme_2030_2050.pdf

5.2 Prüfung der Verfügbarkeit der Daten für die Datenbank

Im AP 5.1 recherchierten und dokumentierten wir die Verfügbarkeit der Daten. Damit verschafften wir uns auch ein Überblick über die Daten, die prinzipiell für eine Ablage in der Datenbank in Frage kämen. Die tatsächliche Verfügbarkeit der Daten prüften wir nun im Arbeitspaket 5.2 aus verschiedenen Sichtweisen:

Verfügbarkeit aus technischer Sicht

Alle untersuchten Szenariendaten lagen in maschinenlesbarem Format vor. Aus technischer Sicht gab es daher keine generellen Verfügbarkeitsprobleme zu lösen. Da die Daten jedoch nicht originär in dem Format vorlagen, wie sie durch die API Schnittstelle verarbeitet werden können, hatte dies Implikationen hinsichtlich der Konversion (Siehe AP 5.3).

- Die Daten des Szenarios KS95 aus der Studie *Klimaschutzszenario 2050* lagen im Excelformat in verschiedenen, unterschiedlich strukturierten Dateien vor.
- Die Daten der Szenarien aus der Studie *Untersuchungen zur Energiestrategie Brandenburgs* lagen im CSV-Format vor und wurden bereits in die OEP geladen und mit Metadaten versehen. Ergänzend wurde ein überarbeitetes Modell der selben Region mit ähnlichem Szenariorahmen namens *oemof-B3 - Analysis of the energy system of Brandenburg and Berlin* erstellt und bereitgestellt.
- Die Daten aus der Studie *Transformationspfade im Wärmesektor: Betriebs- und volkswirtschaftliche Betrachtung der Sektorkopplung mit dem Fokus Fernwärme mit hohen Anteilen konventioneller KWK-Erzeugung und Rückkopplung zum Gesamtenergieversorgungssystem* lagen ebenfalls im Excelformat in verschiedenen, unterschiedlich strukturierten Dateien vor.

Verfügbarkeit aus vertraglicher Sicht Die vertraglich festgelegten Bedingungen für das **KS-95-Szenario** bargen Herausforderungen hinsichtlich der Verfügbarkeit der Daten für die Datenbank. Wir führten hier eine Nachverhandlung bezüglich der Veröffentlichung unter einer offenen Lizenz durch. Diese Nachverhandlung bereiteten wir mit juristischer Unterstützung vor (siehe dazu Arbeitspaket 5.4) und schlossen sie erfolgreich ab. Die Daten konnten unter der Datenlizenz Deutschland – Namensnennung – Version 2.0³⁸ veröffentlicht werden, solange folgender Text angegeben wird: *Diese Studie / Internetseite wird im Auftrag des Bundesministeriums um Umwelt, Naturschutz und nukleare Sicherheit erstellt und die Logos entsprechend verwendet werden..* Der Text ist in den entsprechenden Datensätzen in den Metadaten integriert, sowie das Logo verlinkt.

Die Szenarien die unter der Monitoring Mechanism Regulation berichtet werden, werden auf dem Central Data Repository EIONET³⁹ unter der Lizenz CC-BY-DK-

³⁸<https://www.govdata.de/dl-de/by-2-0>

³⁹https://cdr.eionet.europa.eu/de/eu/mmr/art04-13-14_lcds_pams_projections/projections/

2.5⁴⁰ veröffentlicht. Das erlaubte die Daten der **deutschen Szenarien** unter der MMR auf der OEP verfügbar zu machen. Wir behielten dabei die Lizenz bei und ergänzten die Datentabellen um weitere relevante Spalten, die künftig eine Erweiterung um zusätzliche Informationen ermöglichen (z.B. die Spalte *member_state*). Die vertraglich festgelegten Bedingungen aus der Studie *Untersuchungen zur Energiestrategie Brandenburgs* erlaubten eine vollständige Veröffentlichung aller Eingangs- und Ergebnisdaten unter offenen Lizenzen. Es wurde für die Daten die Copyleft-Lizenz ODbL-1.0⁴¹ gewählt.

Die vertraglich festgelegten Bedingungen aus der Studie *Transformationspfade im Wärmesektor: Betriebs- und volkswirtschaftliche Betrachtung der Sektorkopplung mit dem Fokus Fernwärme mit hohen Anteilen konventioneller KWK-Erzeugung und Rückkopplung zum Gesamtenergieversorgungssystem* stellten kein Hindernis für eine Veröffentlichung dar.

5.3 Implikationen für Datenformate

Bei Antragstellung gingen wir davon aus, dass die Daten sie in sehr unterschiedlichen Formaten vorliegen würden, z.B. auch nicht-maschinenlesbar. Dies war nicht der Fall. Alle Daten lagen maschinenlesbar vor – entweder als CSV- oder XLS-Dateien.

Die Daten des KS95 Szenarios lagen wie unter AP 5.2 beschrieben in unterschiedlich strukturierten Exceldateien vor. Um diese in die Datenbank zu laden, haben wir die Daten in eine datenbankfreundliche csv Struktur übertragen und den oep-client genutzt, den wir entwickelt haben⁴².

Die Daten der Szenarien aus dem Projekt „Transformationspfade im Wärmesektor“ des Fraunhofer IEE lagen ebenfalls wie unter AP 5.2 beschrieben in unterschiedlich strukturierten Excel-Dateien vor. Um diese in die Datenbank zu laden, wurden sie zunächst in eine datenbankfreundliche Struktur mit einheitlichen Spalten umgeschrieben. Daraufhin wurden die Daten so zusammengefasst, dass je Szenario und Kategorie (z.B. Strombilanz, Fernwärmebilanz) die Daten aus allen Szenariojahren (2030 und 2050) zusammengefasst wurden. Zum Hochladen in die Datenbank wurden Python-Skripte auf Basis einer Anleitung des Öko-Instituts erstellt.

Die Daten der RLI-Szenarien lagen bereits als gut strukturierte CSV-Daten vor. Dadurch konnten diese problemlos in die OEP geladen werden.

Da eine explizite Dokumentation von Konversionsnotwendigkeiten der Daten innerhalb und außerhalb der Datenbank für die exemplarischen Datensätze nicht notwendig war, fokussierten wir die Arbeiten im weiteren Verlauf von AP 5.3 um. Wir erstellten und veröffentlichten zwei Briefings, die in Kombination mit konzeptionellen

⁴⁰<https://creativecommons.org/licenses/by/2.5/dk/deed.en>

⁴¹<https://opendatacommons.org/licenses/odbl/1-0/>

⁴²

tionellen Überlegungen in Arbeitspaket 3.3. auch in Handreichungen auf der OEP mündeten (siehe dazu auch Links zu Tabellen in AP 3.3). Die Briefings beziehen sich auf generelle Konversionsnotwendigkeiten in Klima- und Energiemodellierung, sowie auf verschiedene Datenformate. Die Briefings hat das Konsortium gemeinsam geschrieben und unter der CC-BY-4.0 Lizenz auf zenodo veröffentlicht:

- **Conversion Necessities in Climate and Energy System Modelling.**

<https://doi.org/10.5281/zenodo.3949692>

Abstract:

With this briefing we provide definitions for units and metrics commonly used in climate and energy modelling. We provide this briefing, because in climate and energy system modelling, the same data can be represented using a diverse set of units and metrics. This leads to challenges when it comes to comparability and re-use. Only when we are certain that a variable is measured in a specific unit or metric, we can use it without the chance of producing errors. For example we may want to compare results of different projections with each other or apply a specific variable in our own modelling work. Sometimes, a conversion is needed which may not be straight forward. We cover units in Section 3, greenhouse gases in Section 4, monetary values in Section 5 and geographic coordinates in Section 6. Throughout these sections, we provide – where available – links to conversion tables hosted on the Open Energy Platform. Like this, we combine what we learn below with practical and open access tools for everyday work. This briefing may develop further and is open to suggested additions by anyone. Please post suggestions on github. Please attach the label *conversion-briefing* to your issue. This work was supported by grant 03ET4057A-D (SzenarienDB) from the Federal Ministry for Economic Affairs and Energy of Germany (BMWi).

- **File Formats in Climate and Energy System Modelling**

<https://doi.org/10.5281/zenodo.3957445>

Abstract:

With this briefing we provide relevant definitions for file types, formats, data bases and concepts commonly used in climate and energy system modelling. We offer this document, because not all data is created equal. We have to deal with a multitude of data sources for input and output. Some information comes in special formats and some information can be stored in different ways. Each way comes with specific advantages and drawbacks. Sometimes, conversion is needed and may not be straight forward. These particularities are important with respect to reproducibility, comparability, re-use and documentation. We lay out a reference to allow a productive approach in working with commonly used file formats and related issues in climate and energy system modelling.

We introduce binary and text based file formats, the difference between vector and raster formats, the classification of the structuredness in data, and the use of databases. Subsequently, we provide a description of specific commonly used formats from all categories. Whenever possible we reference an external resource for further explanation. Finally, we conclude with a short section of best practices and recommendations. This briefing may develop further and is open to suggested additions by anyone. Please post suggestions here on github. Please attach the label *formats-briefing* to your issue. This work was supported by grant 03ET4057A–D (SzenarienDB) from the Federal Ministry for Economic Affairs and Energy of Germany (BMWi).

5.4 Fragen der Datenlizenzierung

Das Arbeitspaket 5.4 widmete sich den Fragen der Datenlizenzierung. Für die Klärung aller Fragen nahmen wir die Expertise des Berliner Juristen Dr. Till Jaeger in Anspruch. Er ist Fachanwalt für Urheber- und Medienrecht sowie Mitbegründer des Instituts für Rechtsfragen der Freien und Open Source Software (ifrOSS).

- Die Daten des Projekts *Klimaschutzszenario 2050* waren vertraglich nicht ohne Weiteres durch das Öko-Institut veröffentlichbar. Daher holten wir mit Hilfe des Juristen die Genehmigung ein, die Daten unter der Lizenz Datenlizenz Deutschland – Namensnennung – Version 2.0 (dl-de/by-2-0)⁴³ unter Berücksichtigung bestimmter Voraussetzungen zu veröffentlichen.
- Die vertraglich festgelegten Bedingungen aus der Studie *Untersuchungen zur Energiestrategie Brandenburgs* erlaubten einen Upload und offene Lizenzierung unter der Open Data Commons Open Database License 1.0 (ODbL-1.0). Daher erfolgte hier keine gesonderte juristische Begleitung.
- Für das aktualisierte Modell *oemof-B3 – Analysis of the energy system of Brandenburg and Berlin* wurde die Lizenz CC-BY-4.0 gewählt.
- Nach den vertraglich festgelegten Bedingungen aus der Studie *Transformationspfade im Wärmesektor: Betriebs und volkswirtschaftliche Betrachtung der Sektorkopplung mit dem Fokus Fernwärme mit hohen Anteilen konventioneller KWK Erzeugung und Rückkopplung zum Gesamtenergieversorgungssystem* wurden die Daten analog zum KS-95-Szenario des Öko-Instituts mit der offenen Datenlizenz Deutschland – Namensnennung – Version 2.0 (<https://www.govdata.de/dl-de/by-2-0>) versehen. Zusätzlich wurde die Förderung durch das Bundesministerium für Wirtschaft und Energie (inklusive dessen Logos) ausdrücklich in den Metadaten erwähnt.

⁴³<https://www.govdata.de/dl-de/by-2-0>

5.5 Integration verfügbarer Daten in die Datenbank

In einem letzten Schritt pflegten wir die Szenariendaten, die in den vorangegangenen Arbeitsschritten von Arbeitspaket 5 als geeignet identifiziert wurden in die Datenbank ein. Damit entstanden Best-Practice-Beispiele der Verfügbarmachung von Szenariendaten in nutzer*innenfreundlicher Art und Weise. Erfahrungen die wir bei der Integration der Szenariendaten in die Datenbank machten, informierten die anderen Arbeitspakete, insbesondere Arbeitspaket 4. Auch weitere Fehler und Schwierigkeiten, die bei der Integration der Daten auftraten diskutierten wir gemeinsam. Somit fungierte dieses Arbeitspaket als Test der Funktionalitäten der OEP, die über die ursprünglich antizipierten Herausforderungen (unvollständiger Upload, Nachbesserung der Daten / Fehlerkorrektur, fehlende Dimensionen, fehlende Dimensionswerte, Datenkonflikte) hinausgingen.

Unter anderem folgende Herausforderungen konnten wir identifizieren und diskutieren: Tabellenbenennung (z.B. keine Verwendung von Leerzeichen und Großbuchstaben erlaubt), Datenstrukturen, Lizenzwahl, Verbesserungsbedarf im Prozess des Datenuploads, insbesondere beim Daten-Upload sehr vieler Tabellen gleichzeitig, Dokumentationsbedarf der Uploadprozesse, Verbesserungsbedarf im Prozess und der Dokumentation des Datenreviews, z.B. zur Verwendung der korrekten Datenformate bei der Erstellung der Metadaten.

Als abschließendes Ergebnis dieses Arbeitspaketes steht die exemplarische Befüllung der Datenbank mit Szenariendaten, die der Öffentlichkeit kostenfrei unter einer Lizenz mit möglichst hohem Freiheitsgrad zur Verfügung steht. Insgesamt luden wir im Projektzeitraum 154 Tabellen aus zehn Szenarien mit insgesamt 160628 Datenpunkten hoch, ein Überblick liegt in A.5 bereit.

Alle Partnerinstitute mit Szenariendaten trugen individuell dazu bei, so dass die bereitgestellten Daten vielfältige relevante Bereiche abdecken:

Öko Das Öko-Institut kümmerte sich um die Befüllung der Datenbank mit Datensätzen aus dem Projekt *Klimaschutzszenario 2050* und mit Daten aus der europäischen Projektionsberichterstattung.

Die Integration der verfügbaren Datensätze in die Datenbank erfolgte mit den Daten aus der 2. Modellierungsrunde. Die Integration beinhaltete nicht nur die Daten und Metadaten, sondern auch die Szenariofactsheets.

Die Daten lagen im Excelformat vor. Zur Integration in die Datenbank via API-Schnittstelle wurden die Daten *datenbankfreundlich* aufbereitet. Je Datensatz wurde dabei eine Datei angelegt. Für die Szenariendaten aus der 2. Modellierungsrunde waren dies insgesamt 11 Dateien, die jeweils mehrere Blätter an Daten enthielten. Diese wurden mit Hilfe eines selbst entwickelten Python-Scripts (*oep-client*⁴⁴) via API-Schnittstelle in das schema *model draft* geladen. Jedes Blatt aus den Excel-

⁴⁴<https://github.com/wingechr/oep-client>

dateien wurde auf der Plattform zu einer eigenständigen Tabelle. Insgesamt stehen auf der OEP 31 Tabellen aus der Studie Klimaschutzszenario 2050 zur Verfügung. Je Datei legten wir eine JSON-Datei mit Metadaten an, die den entsprechenden Tabellen auf der OEP zugeordnet wurde. Nach Upload der Datentabellen in *model draft* wurden die Daten via github in den Reviewprozess eingestellt (siehe <https://github.com/OpenEnergyPlatform/data-preprocessing/issues/27> und <https://github.com/OpenEnergyPlatform/data-preprocessing/tree/review/ksz2050>). Nach erfolgreichem Review wurden die Datentabellen in das Schema *scenario* verschoben und sind nun dort abrufbar. Neben dem KS-95-Szenario⁴⁵ luden wir auch die Daten des AMS-(2012)-Szenarios⁴⁶ und des KS-80-Szenarios⁴⁷ aus der 2. Modellierungsrunde in die Datenbank. Somit ermöglichten wir, dass die wesentlichen Daten aller in der 2. Modellierungsrunde gerechneten Szenarien in der Datenbank abrufbar sind.

Da die Daten der 3. Modellierungsrunde des Projektes *Klimaschutzszenario 2050* nicht pünktlich genug für Berücksichtigung in *SzenarienDB* zur Verfügung standen.

Da die OEP perspektivisch auch als europäische Plattform fungieren könnte, testeten wir die Integration von schon veröffentlichten Daten aus europäischer Projektionsberichterstattung (Artikel 14, Monitoring Mechanism Regulation⁴⁸). Diese Daten liegen ebenfalls als csv Format vor. Wir nutzten die Daten, die Deutschland unter der MMR in 2019 und 2017 berichtete und integrierten diese in eine Tabelle. Wir ergänzten zu diesem Zweck Spalten (z.B. Member State, submission year, parameter, scenario). Perspektivisch könnte diese Erweiterung auch eine Integration von weiteren Daten anderer Mitgliedstaaten in die selbe Tabelle erlauben. Die Daten luden wir ebenfalls mit dem oep-client in die Datenbank und lösten den entsprechenden Reviewprozess auf GitHub aus⁴⁹. Die Daten sind nun im Schema *scenario*⁵⁰ verfügbar.

RLI Die OEP hat für das RLI in der Entwicklung von Energiesystemmodellen, welche überwiegend mit dem Open Energy Modelling Framework (oemof) erstellt werden, eine zentrale Bedeutung. Sie dient sowohl der Dokumentation und Publikation von Eingangs- und Ergebnisdaten, als auch der kollaborativen Zusammenstellung von offenen Datensätzen.

Im Rahmen des Projekts hat das RLI die Szenarien der Untersuchungen zur Energiestrategie Brandenburgs auf die OEP geladen, welche in einer Studie für die Partei die Grünen in Brandenburg erstellt wurde. Der Review dazu wur-

⁴⁵<https://openenergy-platform.org/dataedit/view/scenario?query=&tags=234>

⁴⁶<https://openenergy-platform.org/dataedit/view/scenario?query=&tags=232>

⁴⁷<https://openenergy-platform.org/dataedit/view/scenario?query=&tags=233>

⁴⁸<https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32013R0525&from=EN#d1e1446-13-1>

⁴⁹<https://github.com/OpenEnergyPlatform/data-preprocessing/issues/58>

⁵⁰https://openenergy-platform.org/dataedit/view/scenario/szenariendb_mmr_de

de intitutsintern, aber öffentlich auf GitHub durchgeführt und infolgedessen angepasst und als Vorlage für die Reviews aller weiteren Uploads verwandt (siehe <https://github.com/OpenEnergyPlatform/data-preprocessing/issues/8>). Zudem hat das RLI die Szenarien von oemof_b3, einer Analyse des Energiesystems der Region Berlin und Brandenburg mit oemof, auf die OEP geladen. Für die drei Szenarien sind jeweils 15 Tabellen, inklusive Metadaten und Factsheets auf der OEP angelegt. Da die Analyse noch erweitert wird, kann sie als ein Testfall für die Veröffentlichung noch laufender Studien ausgewertet werden.

Nicht als Teil eines bestehenden Szenarios, aber als Datengrundlage für künftige Analysen lud das RLI auch die erste offen lizenzierte Sammlung von Turbinen-Kennlinien von Windenergieanlagen (wind_turbine_library)⁵¹ auf die OEP.

Anhand der genannten Datensätze wurden auch die verschiedenen Wege zum Upload getestet. Während die ersten Datensätze noch mit direktem Datenbankzugriff mittels SQL-Befehlen auf die OEDB geladen wurden, kam für oemof_b3 ausschließlich die zwischenzeitlich von der OvGU Magdeburg entwickelte API zum Einsatz und wurde daran gemessen und ausgebessert. Auf die API aufbauend haben wir schließlich ein Tool entwickelt, welches anhand eines ausgefüllten Metadatenstrings die vereinfachte Erstellung von Tabellen in der OEDB und den anschließenden Upload in diese erlaubt (siehe <https://github.com/OpenEnergyPlatform/oem2orm>).

IEE Das Fraunhofer IEE kümmerte sich um die Befüllung der Datenbank mit Datensätzen aus dem Projekt *Transformationspfade im Wärmesektor: Betriebs- und volkswirtschaftliche Betrachtung der Sektorkopplung mit dem Fokus Fernwärme mit hohen Anteilen konventioneller KWK-Erzeugung und Rückkopplung zum Gesamtenergieversorgungssystem*.

Die Integration der verfügbaren Datensätze in die Datenbank erfolgte mit den finalen Daten der Szenariojahre 2030 und 2050. Die Integration beinhaltete auch hier nicht nur die Daten und Metadaten, sondern auch die Szenariofactsheets.

Die Daten lagen im Excelformat vor. Zur Integration in die Datenbank via API-Schnittstelle wurden die Daten datenbankfreundlich aufbereitet. Für jede Kategorie in jedem Szenario wurde dabei eine Datei angelegt. In diesem Fall waren dies insgesamt 70 Excel-Tabellenblätter (10 Blätter mit Input-Daten sowie für alle 4 Szenarien jeweils 15 Blätter mit Output-Daten), die nun auf der OEP zur Verfügung stehen. Sie wurden mit Hilfe eines selbst entwickelten Python-Scripts auf Basis einer vom Öko-Institut erstellten Anleitung via API-Schnittstelle in das schema model draft geladen. Jedes Excel-Tabellenblatt wurde auf der Plattform als eigenständige Tabelle hochgeladen. Je Datei legten wir eine JSON-Datei mit Metadaten an, die den entsprechenden Tabellen auf der OEP zugeordnet wurde. Nach Upload der Datentabellen in model draft wurden die Daten via github in den Reviewprozess eingestellt.

⁵¹https://openenergy-platform.org/dataedit/view/supply/wind_turbine_library

5.5 Integration verfügbarer Daten in die Datenbank

Nach erfolgreichem Review wurden die Datentabellen in das Schema scenario verschoben und sind nun dort abrufbar.

6 Entwicklung von Konzepten und Handlungsempfehlungen zur kontinuierlichen Aktualisierung/Erweiterung der Daten und zum Weiterbetrieb der Datenbank

Hauptverantwortung Text: Fraunhofer IEE

Zur Bearbeitung dieses Arbeitspakets haben wir ein "Konzeptpapier Weiterbetrieb" entwickelt, in dem wir die einzelnen Teilaspekte des Weiterbetriebs und deren Sicherstellung dokumentiert haben. Es dient neben der Dokumentation der Ergebnisse auch der Abstimmung mit dem Projektträger, der in die Entwicklung dieses Arbeitspaketes eingebunden werden sollte.

Das Arbeitspaket ist in vier Unteraufgaben geteilt. Zunächst entwickelten wir Lösungen für die Anforderungen an den technischen Weiterbetrieb vor. Dann analysierten wir „weiche“, d.h. nicht-technische Anforderungen an den Weiterbetrieb und deren Lösungsansätze. Der Aspekt der Finanzierung des Weiterbetriebes stellt einen weiteren wesentlichen Aspekt dar und schließlich gehen wir in Abschnitt ?? auf die strukturellen Grundlagen und die notwendige Flexibilität für den Weiterbetrieb ein.

Das vollständige Konzept für den Weiterbetrieb befindet sich im Anhang A.6.

7 Fazit

Das Ziel des Projekts bestand darin die bestehende OEDB zu einer frei zugänglichen Szenariendatenbank mit grundlegenden forschungsrelevanten Elementen auszubauen. Dafür haben wir systematisch Anforderungen abgeleitet und von der Energie-Community gesammelt. Nach 3 Jahren Arbeit haben wir eine Datenbankinfrastruktur und geeignete Tools für die frei zugängliche Ablage von Energie- und Klimaschutz-Szenariendaten geschaffen. Durch Bestückung der Datenbank mit Szenariendaten haben wir gezeigt, wie Szenarientransparenz verbessert werden kann und damit auch den Weg für künftige Szenarienvergleiche ein Stück weit geebnet. Wir haben Standards geschaffen, z.B. zur FAIREN Metadatendokumentation auf der OEP, API-Funktionalitäten und der Veröffentlichung von Szenariendaten. Die Open Energy Ontology, deren Entwicklung in SzenarienDB begann, stellt einen einheitlichen und wachsenden Sprachraum der Energiesystemmodellierung zur Verfügung und wird bereits in weiteren Projekten vorangetrieben.

Zudem haben wir aktive Vernetzung innerhalb der Modellierungs-Community betrieben. Eine umfangreiche Listung aller Veröffentlichungen und Vorträge findet sich im administrativen Schlussbericht.

8 Ausblick

Die geschaffenen Standards bilden eine Grundlage für bessere Zusammenarbeit und Interoperabilität in der Energiesystemmodellierung. Wir streben eine weitere Einbindung der OEP und OEO in der Energiesystemmodellierungscommunity an. Eine Reihe von Anforderungen an die Datenbank, konnte innerhalb des Projektes noch nicht vollständig umgesetzt werden. Zudem haben wir weitere Anforderungen an die OEP ausmachen können und insbesondere eine gute Nutzbarkeit als kritisch notwendig ausgemacht. Die weiteren Entwicklungen sollten also der Community einen möglichst einfachen Zugang zur Nutzung der OEP und ihren Funktionalitäten bieten. Nach dem Abschluss des Projektes SzenarienDB werden wir die Entwicklungen im Nachfolgeprojekt SIROP weiterführen. Der Fokus von SIROP liegt in der Interoperabilität und Vergleichbarkeit von Energieszenarien. Das BMWi-geförderte Projekt begann am 01.04.2021 und läuft über 3 Jahre.

Literaturverzeichnis

- [1] Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. Dbpedia: A nucleus for a web of open data. In *The semantic web*, pages 722–735. Springer, 2007.
- [2] Mike Cohen. *User Stories Applied: For Agile Software Development*. Addison-Wesley, Bosten, 2004. ISBN 0-321-20568-5.
- [3] Mike Cohen. *Succeeding with Agile: Software Development Using Scrum*. Addison-Wesley, Bosten, 2010. ISBN 0-321-57936-4.
- [4] Ingo Ellersdorfer. *Energiemodelle zu Innovation und moderner Energietechnik: Analyse exogenen und endogenen technischen Fortschritts in der Energiewirtschaft*, volume 24 of *Umwelt- und Ressourcenökonomik*. LIT, Münster, 2007. ISBN 9783825800260.
- [5] Fraunhofer UMSICHT and Fraunhofer IWES. Abschlussbericht metastudie energiespeicher. URL https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwioML0TwvbdAhXO-6QKHSr_AdQQFjAAegQICRAC&url=https%3A%2F%2Fwww.umsicht.fraunhofer.de%2Fcontent%2Fdam%2Fumsicht%2Fde%2Fdokumente%2Fpressemitteilungen%2F2015%2FMetastudie-Energiespeicher-Kurzfassung-web.pdf&usg=AOvVaw1cS_VM5G6u3hGgW-tM9Dq5.
- [6] Michael D. Kaplowitz, Timothy D. Hadlock, and Ralph Levine. A comparison of web and mail survey response rates. *Public Opinion Quarterly*, 68(1):94–101, 2004. ISSN 0894-4393. doi: 10.1093/poq/nfh006.
- [7] Editha Kötter, Ludwig Schneider, Frank Sehnke, Kay Ohnmeiss, and Ramona Schröer. The future electric power system: Impact of power-to-gas by interacting with other renewable energy components. *Journal of Energy Storage*, 5:113–119, 2016. ISSN 2352152X. doi: 10.1016/j.est.2015.11.012.
- [8] Stefan Lechtenböhrer, Andreas Palzer, Thomas Pregger, Hans Christian Gils, Philip Sterchele, Christoph Kost, Lucas Brucker, Tomke Janßen, Christine Krüger, Dietmar Schüwer, et al. Regmex-modellexperimente und -vergleiche zur simulation von wegen zu einer vollständig regenerativen energieverorgung: Schlussbericht. 2018. doi: https://epub.wupperinst.org/frontdoor/deliver/index/docId/7096/file/7096_RegMex.pdf.

- [9] Steve McConnell. *Software estimation: Demystifying the black art*. Safari Books Online. Microsoft Press, Redmond, Washington, 2006. ISBN 9780735605350. URL <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10762188>.
- [10] Ngoc Tu Phan. Evaluation und weiterentwicklung der szenariendatenbank in der open energy platform im bezug auf identifizierte nutzeranforderungen. 2021.
- [11] Klaus Pohl. *Requirements Engineering: Grundlagen, Prinzipien, Techniken*. dpunkt.verlag, Heidelberg, 2 edition, 2008. ISBN 978-3-89864-550-8.
- [12] Klara Reder, Mirjam Stappel, Christian Hofmann, Hannah Förster, Lukas Emele, Ludwig Hülk, and Martin Glauer. Identification of user requirements for an energy scenario database. *International Journal of Sustainable Energy Planning and Management*, pages 95–108, 2020. URL <https://dx.doi.org/10.5278/ijsepm.3327>.
- [13] Gerhard Versteegen, editor. *Anforderungsmanagement*. Springer, Berlin, Heidelberg, New York, 2004. ISBN 3-540-00963-9.
- [14] Bill Wake. Invest in good stories, and smart tasks, 2003. URL <https://xp123.com/articles/invest-in-good-stories-and-smart-tasks/>.
- [15] Jeremy C. Wyatt. When to use web-based surveys. *Journal of the American Medical Informatics Association*, 7(4):426–430, 2000. ISSN 0959-8138. doi: 10.1136/jamia.2000.0070426.

A Anhang

A.1 Anhang nach Kapitel

Tabelle A.1: Questions of the online survey on the topic: users expect/wish of a scenario database.

nr	question	answers
1	Are you working as ...?	<ul style="list-style-type: none"> • programmer • natural scientist • engineer • economist • social scientist • other
2	Information about your institute/ organization/ enterprise (optional)	<ul style="list-style-type: none"> • Research center • Government authorities • NGO • System operator / utility company • Other • Position
3	Are you working with scenarios?	<ul style="list-style-type: none"> • Yes, I am involved in the generation of scenarios and working with scenarios generated by others. • Yes, I am working with scenarios generated by others. • Not yet.
4	Would it be an option for you to provide your own scenarios for 'SzenarienDB'?	<ul style="list-style-type: none"> • Yes, I would provide my own scenarios and publish all assumptions, as far as possible. • Yes, in case this is part of my project and will be financed. • No, this is not an option because of the license. • No, this is not an option for me because of other reasons, which are ...
5	Would it be an option for you to include a database like 'SzenarienDB' in your workflow by using scenarios from it?	<ul style="list-style-type: none"> • Yes, sounds good. • No, using scenarios from 'SzenarienDB' is not an option for me, because ...
6	Would it be an option for you to have a port implemented / implement a port by yourself between your models and 'SzenarienDB', which enables an easy access for further usage?	<ul style="list-style-type: none"> • Definitely. • Yes, in case of... • No. • Explanation:
7	There are several definitions and understandings of what a 'scenario' is, in the context of energy system modelling. Which parameters are part of a scenario in your daily work?	<ul style="list-style-type: none"> • General framing parameters and assumptions (e.g. geographical and temporal scope, ...) • Scenario type (e.g. extreme scenario, objective scenario, ...) • Model input data • Justification/ explanation on assumptions • Modelling parameters • Model output data • Other:
8	Are you using (external) databases in your daily work, such as OpenStreetMap, Eurostat, etc., to download data, e.g. as input to your models?	<ul style="list-style-type: none"> • Yes, in many cases. • Rarely. • Never. • When using a database, it is usually one of these:

Tabelle A.1: Questions of the online survey on the topic: users expect/wish of a scenario database.

nr	question	answers
9	How important are the following database features for you when looking for data/ using a database?	<ul style="list-style-type: none"> • Possibility to filter data while searching • Text search • Preview of data, e.g. as table • Ad-hoc visualization of data, e.g. as diagram • Description of metadata as text • Glossary/ ontology
10	In addition, the following features are important to me when looking for data in the internet / using external databases:	
11	Which type of data provision is the most comfortable for your purposes? (Download interface, formats)	<ul style="list-style-type: none"> • API • table • study • .xlsx • .csv • .json • .xml • .pdf • .nc • .cdf • datapackage • Other:
12	Please rate the following criteria for using a database like 'SzenarienDB' according to your personal opinion and order by priority: No. 1 = 'most important' to No. 6 = 'least important'.	<ul style="list-style-type: none"> • List of references for all datasets • Quality check of new scenario data by database crew • Speed • Easy and intuitive upload of your own scenarios • Possibility of processing data directly in the database • Unit conversion in the database
13	Please value the criteria 'Quality of data', 'Quantity of data' and 'User friendliness' for 'SzenarienDB' by distributing 100 points among the three categories. The more important a category is to you, the more points you assign.	<ul style="list-style-type: none"> • Quality of data • Quantity of data • User friendliness
14	Which level of detail do you need for your data?	<ul style="list-style-type: none"> • High (e.g. hourly time series for one year, spacial resolution in scale of kilometers) • Medium (e.g. typical days, ...) • Low (e.g. aggregated values for countries, years, ...) • Further explanations:
15	Do you know the Open Energy Platform? http://openenergy-platform.org/	<ul style="list-style-type: none"> • Yes, I am quite familiar with it. • Yes, but I don't have many experiences with it. • No.
16	Which properties/ features do you appreciate when using OEP?	

Tabelle A.1: Questions of the online survey on the topic: users expect/wish of a scenario database.

nr	question	answers
17	Which properties are uncomfortable for you when using OEP?	
18	News on the project 'SzenarienDB' are available here https://www.ief.fraunhofer.de/de/projekte/suche/laufende/SzenarienDB.html?cq_ck=1529400732290 or here http://reiner-lemoine-institut.de/szenariendb/ Are you interested in getting information via Email about further actions in the project 'SzenarienDB'?	<ul style="list-style-type: none"> • No, thank you. • Yes, please. Email:

ersteSpalte	Spalte2 subzwei	Spalte3 subdrei	letzteSpalte subdrei
A	1	2	3
B	4	5	6
C	7	8	9

Tabelle A.2: Tabellenunterschrift der Anhangstabelle.

Quelle: Quell des Ursprungs

A.2 Auswertung von Plattformen aus Nutzer*innensicht (AP 1.1)

Scale: 0.0 - 1.0

	APIs	Quantities	Qualities	UN	World Bank Open Data	Eurostat	European Union Open Data	Open Street Map	Teststreifen	Climate Data Center	EUMETSAT	FAOSTAT	OEP	Open Energy Data	regio	European Data Portal	Udata
data																	
license	0.11	0.10	0.12	0.40	0.50	0.90	1.00	1.00	0.65	1.00	0.65	1.00	0.65	1.00	1.00	1.00	1.00
temporal coverage	0.05	0.04	0.03	0.90	0.90	0.75	0.60*	1.00	0.90	0.90	0.60*	0.90	0.60*	0.60*	0.60*	0.60*	0.60*
temporal resolution	0.08	0.05	0.03	1.00	0.50	0.50	0.75	1.00	0.55	1.00	0.85	0.50	0.25	0.50	0.50	0.50	0.50
spatial coverage	0.08	0.05	0.03	1.00	0.50	0.50	0.50	1.00	0.25	1.00	1.00	1.00	0.25	0.70	1.00	0.50	1.00
spatial resolution	0.08	0.05	0.03	0.50	0.75	0.50	0.50	1.00	0.25	0.65	0.90	0.65	0.75	0.55	0.40	0.65	0.65
coverage/resolution/other	0.07	0.06	0.06	0.50	0.60	0.85	0.70	1.00	0.25	0.85	0.75	0.75	0.75	0.55	0.40	0.65	0.75
interface																	
find data																	
search/index	0.04	0.05	0.06	0.50	0.25	0.50	0.60	0.85	0.00	0.00	0.00	0.75	0.65	0.75	0.50	0.50	0.85
glossary	0.04	0.05	0.06	0.50	0.75	1.00	0.50	1.00	0.00	0.00	1.00	0.75	0.75	1.00	0.00	0.90	0.85
data preview	0.01	0.04	0.05	1.00	1.00	1.00	0.00	0.75	0.00	0.00	0.25	1.00	1.00	1.00	0.00	0.50	1.00
on site visualisation	0.01	0.03	0.04	0.35	1.00	1.00	0.25	1.00	0.00	0.00	0.50	1.00	0.25	0.00	0.00	0.70	0.30
data description/meta data	0.04	0.05	0.05	0.45	0.50	0.75	0.75	1.00	0.25	0.65	0.75	0.90	0.75	0.75	0.00	0.75	0.85
retrieve data																	
api?	0.15	0.05	0.01	1.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
download formats	0.03	0.05	0.03	1.00	1.00	1.00	1.00	1.00	0.50	0.85	1.00	1.00	0.25	1.00	1.00	1.00	1.00
download selection/filter	0.04	0.05	0.07	1.00	1.00	0.60	0.70	1.00	0.55	1.00	1.00	1.00	0.25	0.70	0.70	0.35	0.60
download restrictions	0.09	0.07	0.06	0.40	0.85	0.50	1.00	0.60	1.00	0.45	1.00	1.00	1.00	1.00	0.70	0.70	0.75
user-friendliness																	
clean user interface?	0.02	0.05	0.07	1.00	1.00	0.75	0.50	1.00	0.50	0.75	0.65	1.00	0.20	0.75	0.75	0.45	1.00
accessible to non-programmers	0.02	0.05	0.07	1.00	1.00	1.00	0.90	0.90	0.50	0.75	0.80	1.00	0.10	0.80	0.90	0.45	0.90
all information easily available?	0.02	0.04	0.06	1.00	1.00	1.00	0.90	0.75	0.23	0.75	0.65	1.00	0.30	1.00	0.80	0.50	1.00
license	0.03	0.05	0.05	0.75	0.75	0.25	0.85	0.90	0.00	1.00	0.85	0.90	0.80	1.00	1.00	0.85	1.00
user feedback	0.01	0.02	0.03	0.65	1.00	0.10	0.75	0.50	0.30	0.30	0.85	0.85	0.75	0.50	1.00	1.00	1.00
social media conclusion																	
APIs	0.70	0.74	0.72	0.79	0.89	0.89	0.89	0.89	0.24	0.63	0.80	0.88	0.84	0.79	0.74	0.66	0.79
Quantities	0.71	0.76	0.69	0.75	0.86	0.86	0.86	0.86	0.27	0.63	0.77	0.90	0.55	0.79	0.73	0.58	0.80
Qualities	0.70	0.77	0.69	0.74	0.85	0.74	0.85	0.85	0.28	0.63	0.74	0.92	0.52	0.79	0.72	0.56	0.81
Total	0.70	0.76	0.70	0.76	0.87	0.87	0.87	0.87	0.26	0.63	0.77	0.90	0.57	0.79	0.73	0.60	0.80
Ranking	10	8	11	7	2	2	2	2	15	12	6	1	14	5	9	13	4

A.3 Vergleich existierender technischer Ansätze

project	name	link	dbs	openess/licence	metadata	metadata_link
MEDEAS	medeas database	https://medeas.cmima.csic.es/medeas/	MariaDB	yes	XML	https://medeas.cmima.csic.es/medeas/faces/public/viewData.xhtml
REflex	reflex database within the ESA ² Data Warehouse	https://data.esa2.eu	PostgreSQL	tbd in detail but yes if possible (CC BY 4.0)	JSON	under development - interim solution: http://reflex-project.eu/public/date-publication/
SET-Nav	SET-Nav model integration platform	https://ob1.ene.iasa.ac.at/SetnavUI	beta version of ixmp scenario explorer, ORACLE db	internal use only		
REEEM	reeem_db & OEP	internal & http://openenergyplatform.org/	PostgreSQL	no	JSON (oemetadata)	https://github.com/OpenEnergyPlatform/examples/tree/feature/meta_data_v1.4/metadata
Cross-project community	OpenEnergyPlatform (OEP)	http://openenergyplatform.org/	PostgreSQL	only open data licenses	JSON (oemetadata)	https://github.com/OpenEnergyPlatform/examples/tree/feature/meta_data_v1.4/metadata
SIM4NEXUS	SIM4NEXUS Database	ftp://193.164.132.202	File Database	To be opened in the future, after result publication		
EUCALC	Trade-offs and Pathways towards Sustainable and Low-carbon European Societies	http://www.european-calculator.eu/	MySQL (subject to change)	Yes, open source licence to be determined	JSON & xls	Will be Bitbucket. Still being developed (by 05/2019)
SPINE	Open source toolbox for modelling integrated energy systems	(TBA)	File Database with Data Packages	CC BY-SA if possible	JSON	
JRC-IDEES	JRC Integrated Database of the European Energy System	http://data.jrc.ec.europa.eu/database/jrc-10110-10001	MS Excel	open but registration at JRC Research Collaboration Portal required	pdf	https://ec.europa.eu/jrc/en/potencia/jrc-ideas
IPCC SR15	IAMC 1.5°C Scenario Explorer hosted by IIASA	https://data.ene.iasa.ac.at/iamc-1.5c-explorer	ixmp scenario explorer, ORACLE db	derivative of CC-BY 4.0, see https://data.ene.iasa.ac.at/iamc-1.5c-explorer/#/license	downloadable as xls	
SENTINEL	The Open Platform	(TBA)		tbd	tbd	
Open ENTRANCE	The Open Platform	(TBA)		tbd	tbd	

A.4 Erhebungsraster Szenariendaten (AP 5.1)

General information on the scenario

- Scenario designation
 - Scenario name
 - Scenario abbreviation
 - Definition and target
 - Target-driven scenario?
 - Measure-driven scenario?
 - New aspects
 - Sensitivities
 - * Type of sensitivities

General information about the study

- Title
 - Title
 - Subtitle
 - Abbreviation
 - General focus
- Originator
 - Authors
 - Contact person
 - Institution(s)
- Funding institution
 - Name
 - Contact person
 - Type of sponsorship
- Publication
 - Year of Publication
 - Citation
 - Digital Object Identifier (DOI)
- Models used

- Name of model(s)
- Type of model(s)

Publication format and license

- Report
 - Published?
 - License
 - Report format
 - URL
- Data
 - Published?
 - License
 - Data format
 - URL

Analysis space

- Temporal
 - Scenario years
 - Weather year, typical year, typical periods
- Spatial
 - Territory
 - International transport
 - Area outside of territory modelled?
 - Upstream chains outside the territory included?

Resolution

- Resolution of the analysis space
 - Temporal (within a scenario year)
 - Spatial resolution of the territory
- Modelled energy sectors
 - Electricity
 - Heating and cooling

- Transport
- Modelled demand sectors
 - Sector definition based on
 - Modelled energy demand sectors
 - * Energy industries / Transformation sector
 - * Manufacturing industries
 - * Tertiary / Commercial
 - * Private households
 - * Transport
 - * Other
 - Modelled non energy sectors
 - * Agriculture
 - * Land use, land use change and forestry (LULUCF)
 - * Waste
 - * Other
- Fuels
 - Fossil fuels
 - Nuclear fuels
 - Renewable fuels
 - Secondary fuels
 - Synthetic fuels
 - Other Fuels

Modelling results

- Energy flows
 - Primary energy
 - Final energy
 - Other energy flows
- Installed capacities
 - Energy producers
 - Energy transport

- Energy storages
- Energy consumers
- Potentials
 - Theoretical potential
 - Technical potential
 - Economic potential
 - Developable potential
- Environmental effects
 - Greenhouse gas emissions
 - * Carbon dioxide (CO₂)
 - * Methane (CH₄)
 - * Nitrous oxide (N₂O)
 - * Fluorinated gases
 - Hydrofluorocarbons (HFC)
 - Perfluorocarbons (PFC)
 - Sulphur hexafluoride (SF₆)
 - Nitrogen fluoride (NF₃)
 - Air pollutants
 - * Carbon monoxide (CO)
 - * Nitric and nitrogen oxide (NO_x)
 - * Non-methane volatile organic compounds (NMVOC)
 - * Sulfur dioxide (SO₂)
 - * Ammonia (NH₃)
 - * Particulate matter (PM₁₀/PM_{2,5})

Exogenous data (modelling input)

- Population
 - Population
 - Working population
- Economic data
 - Gross domestic product (GDP)

- Gross value added (GVA)
- Other
- Prices
 - Primary energy prices
 - * Coal
 - * Petroleum
 - * Natural gas
 - * Other primary energy prices
 - Secondary energy prices
 - * Electricity
 - * District heat
 - * Liquid fuels
 - * Synthetic fuels
 - * Other secondary energy prices
 - CO₂-Prices
 - * ETS certificates
 - * Other
 - Investment cost
 - * Energy producers
 - * Energy transport
 - * Energy storages
 - * Other

Endogenous data (calculated by model)

- Population
 - Population
 - Working population
- Economic data
 - Gross domestic product (GDP)
 - Gross value added (GVA)
 - Other

- Prices
 - Primary energy prices
 - * Coal
 - * Petroleum
 - * Natural gas
 - * Other
 - Secondary energy prices
 - * Electricity
 - * District heat
 - * Liquid fuels
 - * Synthetic fuels
 - * Other
 - CO2-Prices
 - * ETS certificates
 - * Other
 - Investment cost
 - * Energy producers
 - * Energy transport
 - * Energy storages
 - * Other

A.5 Hochgeladene Tabellen (AP 5.5)

- 10 Szenarien
- 154 Tabellen
- 160628 Datenpunkte
- 4 Studien
- Open Database License (ODbL) v1.0: 7
- dl-de-by-2.0: 101
- Attribution 2.5 Denmark: 1
- Keine 45

Scenario: oemof_abb (RLI)

- abbb_constraints (Lizenz: ODbL-1.0)
- abbb_demand (Lizenz: ODbL-1.0)
- abbb_emob (Lizenz: ODbL-1.0)
- abbb_regions (Lizenz: ODbL-1.0)
- abbb_simulation_parameter (Lizenz: ODbL-1.0)
- abbb_transformer (Lizenz: ODbL-1.0)
- abbb_transmission_capacity (Lizenz: ODbL-1.0)

Scenario: oemof_b3 (RLI)

- oemof_b3_simple_scenario_1_biomass_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_bus (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_ch4_gt (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_electricity_curtailment (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_electricity_demand (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_electricity_demand_profile (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_electricity_liion_battery (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_electricity_shortage (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_electricity_transmission (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_oil_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_other_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_solar_pv (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_solar_pv_profile (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_wind_onshore (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_1_wind_onshore_profile (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_biomass_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_bus (Lizenz: CC-BY-4.0)

- oemof_b3_simple_scenario_2_ch4_gt (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_electricity_curtailment (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_electricity_demand (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_electricity_demand_profile (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_electricity_liion_battery (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_electricity_shortage (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_electricity_transmission (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_oil_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_other_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_solar_pv (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_solar_pv_profile (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_wind_onshore (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_2_wind_onshore_profile (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_biomass_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_bus (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_ch4_gt (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_electricity_curtailment (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_electricity_demand (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_electricity_demand_profile (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_electricity_liion_battery (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_electricity_shortage (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_electricity_transmission (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_oil_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_other_st (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_solar_pv (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_solar_pv_profile (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_wind_onshore (Lizenz: CC-BY-4.0)
- oemof_b3_simple_scenario_3_wind_onshore_profile (Lizenz: CC-BY-4.0)

Scenario: fh_iee_trafo_fw_viel_dez_bio_hohe_sani (IEE)

- fh_iee_trafo_fw_input_biomassepotential (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_cop_zeitreihen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_ee_potentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_emissionen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_fahrzeugpotentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_kuppelleistungen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_laender (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_rahmendaten (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_technologieparameter (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_waermepotentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_emissionen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_endenergie (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_fw_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_fw_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_gasf_kws_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_gebaeude (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_h2_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_h2_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_interkonnektoren (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_speicher (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_strom_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_strom_erzeugung (Lizenz: dl-de-by-2.0)

- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_systemkosten (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_verkehr (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_hohe_sani_zeitreihen (Lizenz: dl-de-by-2.0)

Scenario: fh_iee_trafo_fw_viel_dez_bio_mod_sani (IEE)

- fh_iee_trafo_fw_input_biomassepotential (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_cop_zeitreihen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_ee_potentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_emissionen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_fahrzeugpotentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_kuppelleistungen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_laender (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_rahmendaten (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_technologieparameter (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_waermepotentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_emissionen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_endenergie (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_fw_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_fw_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_gasf_kws_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_gebaeude (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_h2_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_h2_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_interkonnektoren (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_speicher (Lizenz: dl-de-by-2.0)

- fh_iee_trafo_fw_viel_dez_bio_mod_sani_strom_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_strom_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_systemkosten (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_verkehr (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_viel_dez_bio_mod_sani_zeitreihen (Lizenz: dl-de-by-2.0)

Scenario: fh_iee_trafo_fw_wenig_dez_bio_hohe_sani (IEE)

- fh_iee_trafo_fw_input_biomassepotential (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_cop_zeitreihen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_ee_potentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_emissionen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_fahrzeugpotentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_kuppelleistungen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_laender (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_rahmendaten (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_technologieparameter (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_waermepotentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_emissionen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_endenergie (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_fw_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_fw_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_gasf_kws_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_gebaeude (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_h2_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_h2_erzeugung (Lizenz: dl-de-by-2.0)

- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_interkonnektoren (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_speicher (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_strom_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_strom_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_systemkosten (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_verkehr (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_hohe_sani_zeitreihen (Lizenz: dl-de-by-2.0)

Scenario: fh_iee_trafo_fw_wenig_dez_bio_mod_sani (IEE)

- fh_iee_trafo_fw_input_biomassepotential (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_cop_zeitreihen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_ee_potentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_emissionen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_fahrzeugpotentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_kuppelleistungen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_laender (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_rahmendaten (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_technologieparameter (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_input_waermepotentiale (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_emissionen (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_endenergie (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_fw_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_fw_erzeugung (Lizenz: dl-de-by-2.0)

- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_gasf_kws_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_gebaeude (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_h2_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_h2_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_interkonnektoren (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_speicher (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_strom_bilanz (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_strom_erzeugung (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_systemkosten (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_verkehr (Lizenz: dl-de-by-2.0)
- fh_iee_trafo_fw_wenig_dez_bio_mod_sani_zeitreihen (Lizenz: dl-de-by-2.0)

Scenario: KSz-2050-R2 AMS2012 (Öko-Institut)

- ksz2050_r2_ams2012_co2_emissions_industrial_processes (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_electricity_consumption (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_electricity_overview (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_energy_and_climate_targets (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_energy_consumption_other_energy_industries (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_fugitive_emissions (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_ghg_emissions_detailed_data (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_ghg_emissions_from_combustion_by_fuel (Lizenz: dl-de-by-2.0)

- ksz2050_r2_ams2012_ghg_emissions_overview (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_parameters_assumptions (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ams2012_primary_and_final_energy_consumption (Lizenz: dl-de-by-2.0)

Scenario: KSz-2050-R2 KS95 (Öko-Institut)

- ksz2050_r2_ks95_co2_emissions_industrial_processes (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_electricity_consumption (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_electricity_overview (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_energy_and_climate_targets (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_fugitive_emissions (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_ghg_emissions_detailed_data (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_ghg_emissions_from_combustion_by_fuel (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_ghg_emissions_overview (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_parameters_assumptions (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks95_primary_and_final_energy_consumption (Lizenz: dl-de-by-2.0)

Scenario: KSz-2050-R2 KS80 (Öko-Institut)

- ksz2050_r2_ks80_co2_emissions_industrial_processes (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks80_electricity_consumption (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks80_electricity_overview (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks80_energy_and_climate_targets (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks80_fugitive_emissions (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks80_ghg_emissions_detailed_data (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks80_ghg_emissions_from_combustion_by_fuel (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks80_ghg_emissions_overview (Lizenz: dl-de-by-2.0)

- ksz2050_r2_ks80_parameters_assumptions (Lizenz: dl-de-by-2.0)
- ksz2050_r2_ks80_primary_and_final_energy_consumption (Lizenz: dl-de-by-2.0)

Scenario: Emissions-Projections-Germany MMR (Öko-Institut)

- szenariendb_mmr_de (Lizenz: Attribution 2.5 Denmark)

A.6 Konzept für den Weiterbetrieb (AP 6)

Konzept für den Weiterbetrieb der OEP-SzenarienDB

(alphabetisch)

Lukas Emele, Hannah Förster, Martin Glauer, Christian Hofmann,
Ludwig Hülk, Klara Reder, Mirjam Stappel, Christian Winger

23. August 2021
Version 1.0.0

1 Einleitung

Dieses Konzept erstellen wir im Rahmen des Forschungsprojektes *SzenarienDB*: Erstellung von Struktur und unterstützenden Features für die frei zugängliche Ablage von Szenariendaten in der Open Energy Database (Förderkennzeichen 03ET4057A-D). Es dient der Dokumentation der Ergebnisse, die wir im Arbeitspaket 6 (AP 6) des Projektes erarbeiten, sowie der Abstimmung mit dem Projektträger, der in die Entwicklung dieses Arbeitspaketes eingebunden werden soll.

Text zum Inhalt des AP 6 aus der Gesamtvorhabensbeschreibung:

In diesem Arbeitspaket wird erarbeitet, wie ein Weiterbetrieb der Datenbank ausgestaltet sein kann. Hierbei wird auf die Erfahrungen der Projektpartner mit internen und öffentlichen Datenbanken aufgebaut werden und insbesondere sollen an dieser Stelle wichtige Akteure eng eingebunden werden, um verschiedene Perspektiven und Aspekte des Weiterbetriebs angemessen und frühzeitig zu berücksichtigen. Die Entwicklung eines Konzepts für den Weiterbetrieb erfolgt in enger Abstimmung mit dem Zuwendungsgeber bzw. dem Projektträger.

Das Arbeitspaket ist in vier Unteraufgaben geteilt, anhand derer dieses Dokument strukturiert ist. Zunächst stellen wir Lösungen für die Anforderungen an den technischen Weiterbetrieb vor (Abschnitt 2). Dann analysieren wir „weiche“, d.h. nicht-technische Anforderungen an den Weiterbetrieb und deren Lösungsansätze (Abschnitt 3). Der Aspekt der Finanzierung des Weiterbetriebes stellt einen weiteren wesentlichen Aspekt dieses Konzeptes dar, den wir in Abschnitt 4 diskutieren. Schließlich gehen wir in Abschnitt 5 auf die strukturellen Grundlagen und die notwendige Flexibilität für den Weiterbetrieb ein.

Zunächst jedoch noch ein kurzer Überblick über die Begrifflichkeiten der Open Energy Platform und Database und die darum herum entwickelte Open Energy Family, die Kern des Projektes *SzenarienDB* bilden.

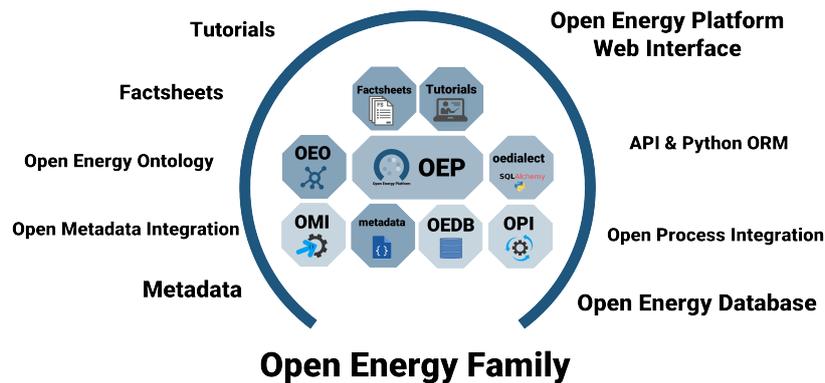


Abbildung 1: Module der Open Energy Family

- *Open Energy Family (OEFamily)*: Ein offenes Framework zum Verarbeiten, Dokumentieren und Veröffentlichenden von Energiesystem-Daten bestehend aus den folgenden Modulen.
- *Open Energy Datenbank (oedb)*: Ein PostgreSQL-Server mit Datenbanken für offene Energiesystem-Daten und plattforminterne Daten (z.B. BenutzerInnen-Verwaltung)
- *Open Energy Platform (OEP)*: Eine mit Django entwickelte Plattform für den Zugriff und die Visualisierung der Daten im Internet.
- *Open Energy Ontology (OEO)*: Eine projektübergreifend entwickelte Ontologie für die Energiesystemanalyse.
- *Open Energy Metadaten (oemetadata)*: Ein für die Energiesystemanalyse entwickelter standardisierter Metadatenstring für die Veröffentlichung von Daten und Datensätzen.
- *Weitere Tools*: Zusätzlich gibt es weitere kleine Module und Packages, die spezifische Aufgaben übernehmen. Z.B. *omi* für die Verarbeitung und Validierung der Metadaten.

Weitere Informationen finden sich hier: <https://openenergy-platform.org/about/>

2 Anforderungen an den technischen und inhaltlichen Weiterbetrieb (AP 6.1)

Aus der Vorhabensbeschreibung des Projektes *SzenarienDB* geht hervor, dass die technischen Anforderungen an den Weiterbetrieb der Datenbank ermittelt werden sollen. Dabei sollen auch grundsätzlich die Varianten „Weiterbetrieb am bisherigen Standort“ sowie „Weiterbetrieb/ Hosting durch Dritte“ untersucht werden. Konkret sollen bei der Analyse der Anforderungen die Aufwände auf Basis der bisherigen Betriebserfahrungen für den Betrieb abgeschätzt werden.

2.1 Auflistung der Anforderungen

Die folgende Liste an Aufwänden wurde aus den im Projekt gemachten Erfahrungen für den Betrieb der Szenariendatenbank und der OEP abgeleitet.

(A) Minimalaufwände für den technischen Weiterbetrieb

- Stromverbrauch
- Rechner, Server
- Netzwerkkosten
- Sicherheitsupdates
- Pflege OEP und OEDB: Einspielen von communityseitigen Aktualisierungen des Plattform-Codes

(B) Aufwände für den inhaltlichen Weiterbetrieb (Personalaufwand, geht über „Minimalbetrieb“ hinaus)

- Aktualisierungen des Plattform-Codes
- Durchführen des Review-Prozess für neue Datensätze
- Hilfestellung, Pflege Tutorials
- Weiterentwicklung OEP
- Weiterentwicklung OEO
- Sonstiges, z.B. Anfragen beantworten

2.2 Prüfung der Varianten für den technischen Weiterbetrieb (A)

Eine Prüfung der beiden Varianten „Weiterbetrieb am bisherigen Standort“ sowie „Weiterbetrieb bzw. Hosting durch Dritte“ wurde durch die Projektpartner

der Universität Magdeburg (OVGU) durchgeführt, die derzeit die Plattform hosten. Der Betrieb am bisherigen Standort hat sich bewährt und ein Standortwechsel wäre im Vergleich zum fortgeführten Betrieb am Standort mit zusätzlichem Aufwand und neuen Unsicherheiten verbunden.

Bei der Wahl eines neuen Standorts müsste die beheimatende Institution, Organisation oder der Anbieter einen langfristigen und ausfallsicheren Weiterbetrieb gewährleisten können. Dazu ist ein gewisses Eigeninteresse am Weiterbetrieb vorteilhaft, gleichzeitig ist den NutzerInnen Unabhängigkeit wichtig, siehe Abschnitt 3.

Für die aktuelle Hosting-Situation ist beides gegeben. Die Universität Magdeburg, vertreten durch Prof. Dr. Till Mossakowski, erklärt sich bereit, die technische Seite des Weiterbetriebs für die nächsten Jahre, mindestens bis zum Jahr 2035, zu gewährleisten und die Bereitstellung für Stromverbräuche, Rechner und Server sowie deren Instandhaltung, Lizenzen und Netzwerkkosten für diesen Zeitraum zu tragen. Hierbei ist es der OVGU möglich, eventuelle Hardwareausfälle im Rahmen der Prioritäten-Liste der Fakultät für Informatik auszugleichen. Damit wäre nicht nur anteilig für den Weiterbetrieb der Szenariendatenbank gesorgt. Das Weiterbestehen der gesamten OEP ist wichtiger Zuverlässigkeitsindikator für die aufgebaute Community. Softwareaktualisierungen auf der Serversystemebene und das Einspielen von communityseitigen Aktualisierungen des Plattform-Codes können vom vorhandenen technischen Personal durchgeführt werden.

Nun stellt sich die Frage, welche Institutionen, Organisationen oder Unternehmen unter diesen Prämissen als externer Host in Frage kämen. Als Anbieter könnte z.B. der Projektträger Jülich selbst fungieren und das Hosting übernehmen. Sowohl Eigeninteresse als auch Unabhängigkeit werden von Projektteam hier als gewährleistet angesehen. Über die technischen und personellen Infrastruktur-Voraussetzungen, die einen nachhaltigen und langfristigen Betrieb zu gewährleisten, können wir als Projektteam jedoch noch keine Aussage treffen, da diese von den internen IT-Strukturen des Projektträgers abhängen. Das plattform- und umgebungsunabhängige Design der OEP und der zugehörigen Datenbank unterstützen einen solchen Übergang.

Generell besteht bei einem Umzug das Risiko, dass die bisherige Stabilität weniger gut gewährleistet werden kann. Die physische Präsenz eines beteiligten Projektpartners an der Hardware sichert einen nachhaltigen Zugriff und die Fortführung der gewählten Architektur. Damit wird eine größere Resilienz erhalten. Diese Vorteile werden als substantiell angesehen und die Variante „Hosting durch Dritte“ wird damit verworfen.

2.3 Anforderungen an den inhaltlichen Weiterbetrieb (B)

Für die inhaltliche Pflege und Weiterentwicklung der Plattform und ihrer Tools fallen vor allem Personalaufwände an. In geringem Rahmen kann und wird bereits die Aktualisierung und Weiterentwicklung der Plattform als Community-

Effort durchgeführt. Dies gilt teilweise auch für die OEO, deren EntwicklerInnen-Community langsam wächst und sich derzeit projektübergreifend etabliert. Als Community-Effort wollen wir langfristig auch den Review-Prozess der Daten und Metadaten auf der OEP etablieren. Hierfür erarbeiten und testen wir derzeit ein Konzept.

Andererseits erfordern die Weiterentwicklung der OEP sowie die Pflege und Anpassung der Tutorials gute Kenntnisse der Plattform sowie gutes Management und Moderation. Neben einem Anteil als Community-Effort kann der inhaltliche Weiterbetrieb am besten als Projektarbeit erfolgen. In Projekten können hierfür zielgerichtet und anwendungsorientiert Zeit und Budget eingeplant werden.

Zur Finanzierung der Personalkosten sind in Abschnitt 4 verschiedene Ansätze erläutert.

Für den rein technischen Weiterbetrieb der OEP (Minimalaufwand) ist der inhaltliche Weiterbetrieb nicht zwingend erforderlich. Ein Weiterbetrieb, der inhaltlich allein auf den Einsatz der Community zählt, birgt jedoch das Risiko, dass die OEP nach einiger Zeit "veraltet" und NutzerInnen wenig anzieht.

3 „Weiche“ Anforderungen an den Weiterbetrieb (AP 6.2)

Welche weichen Anforderungen gibt es aus Sicht der OEP-NutzerInnen für den Weiterbetrieb? Wie werden sie bewertet bzw. priorisiert?

1. Anforderung: Unabhängigkeit des Datenbankbetreibenden

Für die NutzerInnen der OEP ist eine größtmögliche Unabhängigkeit der Datenbankbetreibenden wichtig. Das garantiert, dass NutzerInnen nicht dadurch abgeschreckt werden, dass sie ihre Daten einer potenziell konkurrierenden Institution zur Verfügung stellen müssen.

Lösungsansatz:

Die Datenbank wird momentan und auch weiterhin am Institut für Intelligente Kooperierende Systeme der Universität Magdeburg gehostet, die im Bereich der Energiesystemanalyse kein inhaltlicher Akteur ist. Der Ansatz, alle Daten unter einer offener Lizenz zur Verfügung zu stellen und die offene programmiertechnische Entwicklung der OEP auf GitHub sowie eine umfangreiche Dokumentation, garantiert einen hohen Grad an Transparenz und Unabhängigkeit. Je mehr Daten verschiedener AutorInnen und Institutionen auf der OEP zur Verfügung stehen, desto eher werden neue NutzerInnen bereit sein, ihre eigenen Daten ebenfalls zur Verfügung zu stellen.

2. Anforderung: Gewährleistung des technischen Weiterbetriebs

Als wesentlich für die Nutzung der Plattform wird die Gewährleistung des Weiterbetriebs an sich gesehen, damit sich der Aufwand lohnt. NutzerInnen müssen Zeit investieren: Know-How aufzubauen, Schnittstellen implementieren, Daten hochladen, Daten lizenzieren und die Beteiligung an den Reviewprozessen.

Lösungsansatz:

Den Weiterbetrieb als solchen zu gewährleisten ist in Abschnitt 2 erläutert. Ein Weiterbetrieb, der inhaltlich neben dem Einsatz der Community auch durch Projekte koordiniert und fokussiert wird, ist aus NutzerInnen-Sicht wünschenswert.

3. Anforderung: Geringe Einstiegshürde

Die Einstiegshürde für die Nutzung darf nicht zu hoch, um neue und potenzielle NutzerInnen nicht abzuschrecken.

Lösungsansatz:

Arbeitspaket 1 und Arbeitspaket 4 im Projekt *SzenarienDB* widmen sich diesem Thema. Eine Entwicklung von Materialien, die potentielle MultiplikatorInnen dazu in die Lage versetzen, Einstiegshürden bei neuen NutzerInnen abzubauen und eine selbstorganisierende Verankerung in der Community zu schaffen sind ein weiterer Ansatz. Im Antrag eines Folgeprojektes *SIROP* ist dies Thema des Arbeitspaketes 3.

4. Anforderung: Verankerung in der Community

Eine weitere wichtige weiche Anforderung ist die Verankerung in der Community. Je mehr die Plattform und die vorhandenen Tools genutzt werden und je mehr (gute) Daten offen zur Verfügung stehen, desto mehr ist die Plattform für NutzerInnen interessant.

Lösungsansatz:

Durch Veröffentlichungen, Poster-Beiträge, Verteilermeldungen etc. haben wir im Projektverlauf stetig auf die Existenz und die Weiterentwicklungen der Szenariendatenbank, OEP und OEO aufmerksam gemacht. Die projektübergreifende Arbeit und die Einsetzung eines Steering Committees für die OEO tragen hier zu deren Unabhängigkeit und Vertretung bei. Auch der oben genannte Abbau von Einstiegshürden wirkt positiv auf eine Verankerung in der Community aus. Ein wichtiger erzielter Erfolg für die Attraktivität der OEP ist die Einwilligung der Übertragungsnetzbetreiber, ihre NEP-Szenarien auf der Szenariendatenbank zu veröffentlichen. Zum einen steigert dies die Relevanz der OEP als Datenplattform. Zum anderen stellen die Übertragungsnetzbetreiber damit ein Positivbeispiel für gute Veröffentlichungspraxis und Datentransparenz dar, das weitere NutzerInnen ebenfalls zur Veröffentlichung animieren kann.

4 Finanzierung des Weiterbetriebs (AP 6.3)

Zentraler Aspekt für den Weiterbetrieb der Plattform und Datenbank ist eine Finanzierung über die Projektlaufzeit hinaus. Generell kann für die Nutzung aufgrund des Open-Science-Charakters kein direktes Nutzungsentgelt erhoben werden. Zwar sind Konzepte einer freiwilligen Abgabe z.B. bei Nutzung durch Wirtschaftsunternehmen vorstellbar, jedoch können wir aktuell nicht abschätzen, inwieweit dies die Aufwendungen hinreichend tragen würde. Es werden daher andere Ansätze geprüft und im Folgenden diskutiert.

4.1 Beantragung eines Folgeprojekts

Folgeprojekte können ein Standbein der Weiterentwicklung der Open Energy Family sein, bei dem Unterhaltung und Pflege der Infrastruktur sowie Weiterentwicklungen inbegriffen sind. Es wurden bereits verschiedene Projektanträge/Skizzen eingereicht (Open_Emily, Open_D3), in denen Budget für die Finanzierung hinsichtlich verschiedener Weiterentwicklungen berücksichtigt wurde. Diese wurden abgelehnt. Die Bewertung der Skizze für ein konkretes Folgeprojekt zu *SzenarienDB* mit dem Akronym *SIROP* war positiv. Wir entwickeln momentan den Vollartrag. In *SIROP* wird Budget für die Unterhaltung und Pflege, die (Weiter-)Entwicklung von Tools und für Wissensvermittlung zur Verringerung der Einstiegshürden in der Community eingeplant.

4.2 Gründung eines gemeinnützigen Vereins

Um die Open Energy Family unabhängig von der Projektlage zu machen, kam darüber hinaus die Idee auf, einen Verein zu gründen, der sich um die Finanzierung der Minimalanforderungen (siehe Abschnitt 2) der Plattform und ihrer Tools bemüht, beispielsweise aus Mitgliedsbeiträgen und Spenden.

Wesentlich ist hierbei die Gemeinnützigkeit. Bei Gründung des Vereins muss die Satzung den Vorgaben für Gemeinnützigkeit entsprechen (geregelt in Abgabenordnung und Einkommensteuerrecht). Es erfolgt eine Anmeldung beim Vereinsregister und die steuerliche Anmeldung beim Finanzamt mit Einreichung der Satzung und Antrag auf vorläufige Bescheinigung der Gemeinnützigkeit. Die Gemeinnützigkeit gilt unter Vorbehalt bis zur Einreichung der Steuererklärungen.

Diese Konstruktion ist bei Software-Projekten attraktiv, jedoch noch nicht sehr verbreitet. So wird sie beispielsweise auch für das offene Energiesystem-Framework *oemof* seit einiger Zeit diskutiert. Ein Verein bringt insbesondere urheberrechtliche Vorteile, weil er als juristische Person die Wahrung der offenen Lizenzen garantieren kann. Die Attribution vereinfacht sich. Die Organisation kann unabhängig von den individuellen Beitragenden agieren. Es gilt jedoch zu bedenken, dass freie Software nicht als Grund für Gemeinnützigkeit anerkannt ist und darauf geachtet werden sollte, Wissenschaft und Forschung als Ziele des potenziellen Vereins festzuschreiben.

Momentan ist die OEP an der OVGU Magdeburg angesiedelt. Diese genießt als wissenschaftliche Einrichtung verschiedene Privilegien im Umgang mit Daten. Hierzu liefert die OVGU Informationen zu bisherigen Erfahrungen aus anderen Open-Projekten an der Fakultät. Diese Privilegien sollten unter allen Umständen erhalten bleiben und müssten daher der Satzung entsprechend verankert werden.

Mit dem Betrieb eines gemeinnützigen Vereins ist teilweise ein hoher bürokrati-

tischer Aufwand verbunden, wie z.B. Steuererklärungen, Tätigkeitsberichte, Mitgliederversammlungen etc.

Durch eine Vereinsgründung könnte eine von Projekten unabhängige Konstanz in den Erhalt der Open Energy Family gebracht werden. Die Gründung eines gemeinnützigen Vereins wird generell als gute Option gesehen. Auf Grund des oben genannten organisatorischen Aufwandes wird dies jedoch nicht Teil des Projektes *SzeanrienDB* sein, sondern stellt eine längerfristige Perspektive dar.

4.3 Verankerung der Nutzung in Zuwendungsprojekten (BMWf)

Um den Aufwand für die *Open Energy Family* zu rechtfertigen, ist die aktive Nutzung der Plattform durch Abruf und Bereitstellung von Daten in der Community wesentlich. Die OEP als attraktive und effiziente Option für die Bereitstellung offener Daten und Szenarien wurde dafür entwickelt, die Umsetzung der Open-Science-Prinzipien zu erleichtern. Daher soll das Bewusstsein für die Bedeutung von Open Science seitens ZuwendungsgeberInnen und ForscherInnen gestärkt werden.

Eine Möglichkeit hierzu wäre, die Veröffentlichung von Forschungsdaten unter offener Lizenz bei der Beantragung von öffentlichen Projekten aktiv zu fördern oder sogar einzufordern.

Prinzipiell bildet die Einräumung von Budget für die Bereitstellung von Daten auf der Plattform, d.h. für Schnittstellenimplementierung, Datenaufbereitung, Datenbeschreibung mittels Metadaten, Lizenzierung der Daten in Projektanträgen seitens des PtJ den wesentlichen Grundstein. Wünschenswert wäre darüber hinaus, die Nutzung durch eine oder mehrere der folgenden Maßnahmen in Projektanträgen zu fördern:

- Vorschlag/Ermutung zur Veröffentlichung der Forschungsergebnisse und Datensätze unter geeigneten offenen Lizenzen in Projektanträgen seitens des PtJ, z.B. unter Einbeziehung bestehender freier Software und offener Plattformen.
- Bevorzugung von Projektanträgen seitens des PtJ bei einer geplanten Veröffentlichung der Forschungsergebnisse und Datensätze unter offener Lizenz.
- Forderung der Veröffentlichung der Forschungsergebnisse und Datensätze unter offenen Lizenzen seitens des PtJ. Überprüfung der Auswahl und Durchführung der Lizenzierung bei Projektabschluss (z.B. im Erfolgskontrollbericht)

Für AntragstellerInnen kann dabei der Hinweis darauf hilfreich sein, dass das eigenständige Hosten von Daten, insbesondere für die Veröffentlichung für Dritte, mit nicht zu unterschätzendem Aufwand und Risiko verbunden ist. Platt-

formen wie die OEP sind eine wertvolle Dienstleistung, denn sie reduzieren Bürokratie und Risiken.

Für welche Zuwendungsprojekte die Veröffentlichung von Forschungsdaten unter offener Lizenz möglich und sinnvoll ist, muss erarbeitet werden. Soll sich dies auf alle Projekte beziehen, auf Referenzprojekte, oder auf Projekte, welche einen Bezug zu Open Source, Open Data oder Open Science nehmen? Im Bereich der Energiesystemmodellierung ist hier mit der Open Energy Family bereits eine gute Infrastruktur verfügbar, sodass die Veröffentlichung niederschwellig ist. Eine Art Leitfaden „How to publish Open Data in öffentlichen Forschungsprojekten“ könnte hier helfen, relevante Informationen zugänglich zu machen. Eine erste Version zu offenen Datenlizenzen wurde bereits auf der OEP veröffentlicht.¹

4.4 Verankerung der Nutzung in Auftragsforschung für Ministerien

Ergebnisse aus Auftragsforschungen für Ministerien sind ebenfalls für die OEP interessant. Zum einen können diese Projekte Daten der OEP nutzen und damit Auftragsforschung effizienter machen. Zum anderen können Zuwendungsprojekte davon profitieren, wenn auch Auftragsforschungsergebnisse offen lizenziert und transparent bereit gestellt werden. In der Regel werden Ergebnisse aus Auftragsforschung für Ministerien in Berichtsform veröffentlicht - wenn bisher auch nicht immer mit maschinenlesbaren Daten und offen lizenziert. Eine systematische Bekanntmachung der OEP und ihrer Vorteile für Auftragsforschung könnte daher den NutzerInnenkreis der OEP vergrößern und die Notwendigkeit des Weiterbetriebs erhöhen. Dazu gehört es auch, darzulegen, warum bzw. wie eine offene Lizenzierung keine Nachteile für die Auftragsforschung mit sich bringt. *So offen wie möglich, so geschlossen wie nötig* könnte dabei einen Leitsatz darstellen, der auch AuftragsforscherInnen von der Nutzung der OEP überzeugen könnte.

Das Öko-Institut hat mit dem ersten offen lizenzierten Datensatz eines Auftragsforschungsprojektes auf der OEP gezeigt, dass auch Ergebnisse der Auftragsforschung veröffentlicht werden können.² Auf diese Erfahrungen kann aufgebaut werden.

Ein Dialogprozess zwischen entsprechenden Organisationen (ZuwendungsgeberInnen, AuftraggeberInnen, AuftragnehmerInnen mit Erfahrung der offenen Lizenzierung) könnte eine Strategie darstellen, um die Notwendigkeit des Weiterbetriebs der OEP auf noch stärkere Beine zu stellen. In diesem Prozess kann den auftraggebenden Institutionen dargelegt werden, dass die Einstellung von Projektmitteln für die Interaktion mit der OEP vorteilhaft gegenüber geschlosseneren Ansätzen ist. AuftragnehmerInnen können anhand von Positivbeispielen Bedenken genommen werden.

¹https://github.com/OpenEnergyPlatform/tutorial/blob/master/other/tutorial_open-data-licenses.ipynb

²<https://openenergy-platform.org/dataedit/view/scenario?query=&tags=185>

5 Strukturelle Grundlagen und Flexibilität

Zu den Anforderungen an einen langfristigen und sicheren Betrieb der Datenbank und Plattform zählt auch die Prüfung und Schaffung der notwendigen struktureller Grundlagen. Hierzu zählen eine saubere Dokumentation der Datenbankstruktur, die primär in AP 3 angelegt wird, sowie das Anlegen einer gewissen Flexibilität im Aufbau, insbesondere im gesamten Programmcode.

5.1 Saubere Dokumentation

Die gesamte programmiertechnische Entwicklungsarbeit und die Diskussionen der OEP-ProgrammiererInnen dazu werden fortlaufend auf GitHub dokumentiert. Für jedes Release wird ein Projekt angelegt, in dem die geplanten und letztendlich durchgeführten Arbeiten zusammengefasst sind.

In projektinternen Redmine-Tickets werden die geplanten Arbeiten gesammelt und für den Projektverlauf, sowie für Zwischen- und Abschluss-Berichte dokumentiert.

Im Zuge der Umstrukturierung und des Redesigns der OEP sowie der Erstellung neuer, aktueller Tutorials wurde die Dokumentation aktualisiert und erweitert. Hier muss ggf. noch ein Prozess für die Dokumentationskontinuität erstellt werden.

5.2 Flexibilität

Die OEP baut auf dem Webframework Django auf, welches in Python geschrieben ist. Die Wahl der weit verbreiteten Architektur ermöglicht theoretisch einen leichten Serverumzug. Durch die hohe Verbreitung wird im Gegensatz zu einer potentiellen Eigenentwicklung auch eine breite Basis an EntwicklerInnen angesprochen, die leicht in die Weiterentwicklung einsteigen können. Viele Standardaufgaben können durch bestehende und gepflegte Module schnell implementiert werden und durch die Wahl von Python als Programmiersprache existiert ein großes Ökosystem von zusätzlichen Modulen. Die im Hintergrund laufende PostgreSQL-Datenbank, welche die Basis für die OEDB bildet, ist ebenfalls sehr weit verbreitet (insbesondere in der Forschung) und wird fortlaufend gepflegt. Der Wechsel zu einer anderen Datenbanksoftware oder auch zu einem anderen Webframework wäre schwierig und brächte keine Vorteile mit sich.

6 Fazit

Auch über das Projektende von *SzenarienDB* hinaus erfolgt das Hosting der Datenbank und OEP wie bisher an der Universität Magdeburg. Die für den Weiterbetrieb notwendigen Minimalanforderungen können hier personell und finanziell abgedeckt werden.

Für einen Weiterbetrieb, der die Minimalanforderungen übersteigt und beispielsweise auch Weiterentwicklungen an der Datenbank und -plattform beinhaltet,

ist eine Finanzierung durch Folgeprojekte notwendig. Ein entsprechender Antrag wird derzeit vorbereitet. Als langfristige Option wird die Gründung eines gemeinnützigen Vereins als Träger der Datenbank erwogen. Hierzu wird jedoch aktuell noch kein Handlungsbedarf gesehen.

Die strukturellen Grundlagen an Flexibilität und Dokumentation wurden auf ein Niveau gebracht, auf dem etwaige personelle oder strukturelle Änderungen bei Instandhaltung und Weiterentwicklungen möglich sind.

Zwei wesentliche Grundsteine für den Weiterbetrieb werden darüber hinaus in der Verankerung der OEP in der NutzerInnen-Community, sowie in der zunehmenden Gewichtung von Open-Science-Prinzipien in öffentlich geförderten Projekten gesehen. Beides wurde im Projekt *SzenarienDB* angestoßen und vorangetrieben. Dennoch ist es unerlässlich für den langfristigen Erfolg der Plattform, diese Entwicklungen auch nach Projektende weiterzuführen.