

Integration eines digitalen Maschinenparks in ein Forschungsdatenmanagementsystem

Osman Altun^{1*}, Tatyana Sheveleva², André Castro², Pooya Oladazimi², Oliver Koepler², Iryna Mozgova¹, Roland Lachmayer¹, Sören Auer²

¹ Institut für Produktentwicklung und Gerätebau, Leibniz Universität Hannover

² TIB - Leibniz-Informationszentrum Technik und Naturwissenschaften und Universitätsbibliothek

* Korrespondierender Autor:

Osman Altun
Institut für Produktentwicklung Gerätebau
An der Universität 1 (8143)
30823 Garbsen
Telefon: +49 511 762 3333
Mail: altun@ipeg.uni-hannover.de

Abstract

The global trend towards comprehensive digitization of technologies in product manufacturing leadings to radical changes in engineering processes and requires a new, expanded understanding in the handling of data. Especially in large interdisciplinary projects with several subprojects, the use of a research data management (RDM) system is necessary.

This paper describes the concept to realise a FDM system according to FAIR (Findable, Accessible, Interoperable, Reusable) data principles and using open source systems. The approach is explained on the example of a digital machine park within the Collaborative Research Centre Oxygen-free production (CRC 1368).

Keywords

Research data management, knowledge management, semantic information linking, FAIR data principles, digitization of scientific data

1. Einleitung

Das effiziente Management von Daten und Informationen, die in Forschungsprozessen anfallen und die Gewinnung von Wissen daraus sind kritischer Erfolgsfaktoren für große Verbundprojekte [1]. Forschungsdatenmanagementprozesse sind Ingenieurwissenschaften aktuell noch stark fragmentiert [2]. In vielen Fällen gibt es keine einheitlichen Standards für die Beschreibung der Daten und die Dokumentation der Datenerzeugung. Dabei ermöglicht die umfassende Dokumentation und zeitnahe Bereitstellung von Daten über mehrere zusammenhängende Teilprojekte hinweg allen Beteiligten eine effizientere Verarbeitung und Analyse. Daten Repositorien dienen als Forschungsdatenmanagementsystem (FDM) der Archivierung und dem späteren Zugriff auf Daten, sie werden durch Metadaten umfangreich beschrieben. Wissensmanagementsysteme (WMS) ergänzen die Dokumentation der Daten-Provenienz: Warum, wie, wo, wann und von wem wurden Daten erzeugt. Die Verknüpfung und kombinierte Nutzung solcher Systeme erzeugt Mehrwerte für Forschende, insbesondere bei der Identifizierung von Zusammenhängen. Amorim, R. C. et al. beschreiben in [3] bekannte FDM-Systeme, die von Institutionen eingesetzt werden können, um den FDM-Workflow zu unterstützen. Sie sind in großen Verbundprojekten wichtige Hilfsmittel bei der Implementierung der FAIR-Data-Prinzipien (*Findable, Accessible, Interoperable, Reusable*) [4]. Diese Prinzipien beschreiben unter anderem die Anforderungen, um Daten mit detaillierten Metadaten zu versehen und so Informationen über den Kontext eines Datensatzes zu liefern und die korrekte Interpretation dessen zu gewährleisten. Bei der Entwicklung von FDM-Systemen sind neben den FAIR-Data-Prinzipien auch weitere Ziele zu berücksichtigen. Eine allgemeine Beschreibung der Ziele von FDM-Systemen ist in [5] zusammengefasst.

Die digitale Darstellung von Materialien, Arbeitsprozessen, Geräten und Betriebsmitteln ist ein wichtiger Aspekt der gemeinsamen Arbeitsorganisation in einem großen Verbundprojekt und trägt dazu bei, aus Daten Informationen und Wissen zu generieren. Dazu muss eine Reihe von Fragestellungen berücksichtigt werden:

- Wie sind intuitive FDM-Werkzeuge zur Unterstützung von Forschungsdatenverarbeitungsprozessen zwischen mehreren Projekten zu realisieren?
- Wie lässt sich ein domänenspezifisches Vokabular zum semantischen Annotieren von Forschungsdaten entwickeln?
- Wie können Daten in einem FDM-System untereinander und mit einer Wissensbasis vernetzt werden und daraus neue Erkenntnisse und Forschungsfragen abgeleitet werden?

Dieser Beitrag beschreibt das Konzept eines semantisch vernetzten Systems aus Daten-Repositorium und Wissensmanagementsystem zur digitalen Darstellung und Verknüpfung von Forschungsobjekten, Maschinen, Betriebsmitteln, Prozesse und Datenflüssen über Teilprojekte eines Verbundprojekts. Es wird ein Überblick über aktuelle Daten-Repositorien, Ansätze zum Umgang mit Metadaten und die Erstellung von domänenspezifischen kontrollierten Vokabularen und Ontologien gegeben. Das allgemeine Konzept des Systems wird beschrieben und die Integration eines digitalen Maschinenparks in das System gezeigt.

2. Stand der Forschung

2.1. Forschungsdaten-Repositorien

Im internationalen Repository-Verzeichnis re3data.org werden über 589 Repositorien mit Bezug zu den Ingenieurwissenschaften gelistet, wobei 81 davon in Deutschland verwaltet werden. Die dort verzeichneten Repositorien weisen eine große Heterogenität bzgl. verwendeter Workflows für die Datenarchivierung und -publikation, (Meta-)Datenstandards und eingesetzter Persistenter Identifiersysteme (PID) auf. Auffällig ist, dass v. a. keine

fachspezifischen Metadaten aus dem ingenieurwissenschaftlichen Bereich bedient werden, sondern die veröffentlichten Datensätze vielmehr auf Metadaten-Standards anderer Disziplinen, z. B. *Data Cube Vocabulary*, *CIF*, *netCDF*, oder auf generische Metadaten-Standards, z. B. *Dublin Core (DC)*, *Data Catalogue Vocabulary (DCAT)*, *DataCite*, *ISO 19115*, *DDI* gemappt werden. Die Nutzung generischer Metadaten schränkt die Auffindbarkeit fachspezifischer Forschungsdaten deutlich ein. Die Spezifizierung der Suche nach den betriebsmittelrelevanten Inhalten im Verzeichnis von re3data.org liefert nur 3 relevante, darunter das SENSOR (<https://sensor.awi.de/>), das Repository zur Aufnahme von Metadaten zu Forschungsplattformen, Geräten und Sensoren, die in Rahmen der Erhebung von Forschungsdaten verwendet werden [6].

2.2. Metadaten, kontrollierte Vokabulare, Klassifikationssysteme und Ontologien

Zum Beschreiben und Bereitstellen von Inhalten in einem FDM-System werden Verzeichnisse, Thesauri, kontrollierte Vokabulare und Klassifizierer verwendet. Kontrollierte Vokabulare und Klassifikationen sind Dokumentationssprachen, die zur inhaltlichen Beschreibung von Objekten, z. B. Forschungsdaten, verwendet werden. Sie werten Metadaten wesentlich auf, ebenso wie die Nutzung eines standardisierten Schemas. Ein Thesaurus ist eine natürlich-sprachliche, geordnete Sammlung von Begriffen und deren Beziehungen zueinander. Die Vergabe von kontrollierten Schlagworten aus einem Thesaurus und die Eingruppierung in bestimmte Bereiche erleichtern die Auffindbarkeit der Daten. Eine Klassifikation dient der Zuordnung von Objekten in Klassen, die durch bestimmte Merkmale charakterisiert sind. Sie beziehen sich auch auf hinterlegte Produktdaten, beispielsweise auf einen Festigkeitswert des gewählten Materials [7]. Das in Ingenieurwissenschaften bereits etablierte Klassifikationssystem ist die DIN 4000-1, die auf Basis einer hierarchischen Struktur die Komponenten-Eigenschaften wie Masse, Farbe, Dicke oder Gesamtlänge erfasst [7]. Die DIN 4000 wurde bereits zum Bestandteil des digitalisierten ISO/IEC-konformen Stammdaten-Standards für Produkte und Dienstleistungen ECLASS, der aktuell über eindeutige maschineninterpretierbare Identifier für über 45.000 Produktklassen und 19.000 Produktmerkmale verfügt (<https://www.eclass.eu/>). Die Wiederverwendung dieser Identifier wird jedoch durch die kommerziell ausgerichtete Nutzungslizenz gehindert. Die semantische Erschließung von Forschungsdaten beruht im Wesentlichen auf Technologien, die in der W3C Semantic Web Activity entwickelt und als *W3C Recommendations* veröffentlicht wurden. Das Semantic Web beschreibt die Erweiterung des *Web of Documents* zu einem *Web of Data*. Im *Web of Data* sind Entitäten zu Linked-Data-Graphen verknüpft. Daten und Informationen unterschiedlichster Herkunft können so auf Basis von Standards maschinell interpretiert, verarbeitet und verknüpft werden [8]. Als ausdrucksstarkes Datenmodell wird das *Resource Description Framework (RDF)* verwendet, das bereits 2004 als *W3C Recommendation* veröffentlicht wurde. Hierbei werden Ressourcen über Beziehungen maschinenlesbar verknüpft. Dazu ist ein Vokabular (maschinenlesbare Spezifikation) fachspezifischer Beziehungen und Klassen notwendig. Basierend auf den semantischen Technologien kann das für das Web-of-Data im Jahr 2007 gestartete W3C Gemeinschaftsprojekt Linked Open Data (LOD) (lod-cloud.net) als Initialzündung angesehen werden. Die Daten im Web-of-Data werden mithilfe gemeinsam genutzter Vokabulare zusammengehalten. Das Linked Open Vocabulary (LOV) Portal [9] enthält momentan 757 Vokabulare. Hierzu gehören z. B. das Schema.org (schema.org).

2.3. Datenmanagementsysteme und Wissensmanagementsysteme

Amorim, R. C. et al. beschreiben in [3] bekannte FDM-Systeme, die von Institutionen eingesetzt werden können, um den FDM-Workflow zu unterstützen. Eines davon ist das *Comprehensive Knowledge Archive Network (CKAN)*, das von der CKAN Association, einem

Teil der Open Knowledge Foundation, entwickelt wurde. CKAN findet als Werkzeug zur Realisierung von offenen Verwaltungsdatenportalen Anwendung und ist inzwischen auf über 1.000 Datenportalen weltweit im Einsatz, wie Data.gov und Data.gov.uk. CKAN findet zunehmend auch Anwendung für Forschungsdaten, wie beispielsweise im Leibniz Data Manager (datamanager.tib.eu) oder dem European Data Portal (europeandataportal.eu) demonstriert wird. CKAN verfügt über offen-konfigurierbare Schnittstellen (APIs) und die Erweiterbarkeit des Dienstes über Plugins und (fachspezifische) Vokabulare. So kann der Dienst um relevante Tools, die z. B. die Ausführbarkeit von Softwarecodes, z. B. Jupyter-Notebooks, oder die Visualisierung von Forschungsdaten erlauben, z. B. AutoCAD .dwg files, ergänzt werden. Mit der Visualisierungsfunktion können z. B. Datensätze auf ihre Relevanz für das eigene Forschungsgebiet geprüft werden (eine sogenannte Screening-Funktion), ohne die Daten vorher zeitintensiv auf den Rechner des Wissenschaftlers herunterzuladen.

Für die Erfassung, Speicherung und Verfügbarmachung semistrukturierter Inhalte haben sich Wikis etabliert. Wikipedia ist das bekannte Beispiel und wird mit der Open-Source-Software MediaWiki betrieben. Wikis haben sich mittlerweile auch in Firmen und in der Forschung als Wissensmanagementsystem zur Dokumentation von semistrukturierter Forschungsinformation, z. B. Protokoll- und Prozessbeschreibungen, etabliert. Für eine erhöhte Durchsuchbarkeit und semantische Annotation der Inhalte wurde die MediaWiki Erweiterung Semantic MediaWiki entwickelt. Semantische Erweiterung ermöglicht die Erfassung und Durchsuchung semistrukturierter Informationen anhand eines maschinenlesbaren Vokabulars, das systemübergreifend die terminologische Grundlage bietet und somit die Dateninteroperabilität fördert.

3. Forschungsdatenmanagementsystem in einem Verbundprojekt

Im Sonderforschungsbereich (SFB) 1368 Sauerstofffreie Produktion arbeiten rund 40 Forscherinnen und Forscher unterschiedlicher Fachrichtungen in mehreren Teilprojekten organisiert gemeinsam an den Grundlagen auf dem Weg zu sauerstofffreien Produktionsprozessen. Im Rahmen dieses Verbundprojektes werden Prozesse und Wirkzonen in sauerstofffreier Atmosphäre zur Entwicklung zukunftsfähiger Produktionstechnologien erforscht [10; 11]. Die Heterogenität von Forschungsdaten reicht dabei von Messdaten, Daten aus Simulationen bis hin zu Bildern und Videos. Auffindbarkeit, Zugänglichkeit, Verständlichkeit und Wiederverwendbarkeit von Daten spielen eine entscheidende Rolle in großen Verbundprojekten, in denen Daten zwischen Projekten oder entlang einer Prozesskette ausgetauscht werden [12]. Im Hinblick auf die Beantwortung in Kapitel 1 formulierter Fragestellungen sollen die Gesamtheit aller Prozessparameter, Metadaten und verwendete Hard- und Software umfassend erfasst, dokumentiert und nutzbar gemacht werden. Methoden und Werkzeuge zur systematischen Speicherung, semantischen Repräsentation, Verarbeitung und Analyse von Informationen in Wissenschaftsprozessen der Produktionsverfahren sollen angewandt werden, die die Qualität des Forschungsdatenmanagements bei vertretbarem Zeit- und Ressourcen-Aufwand sicherstellen. Zur Entwicklung eines nachhaltigen FDM-Systems sind in Umfragen mit den Forschenden der einzelnen Teilprojekte sämtliche Forschungsaktivitäten, Datenstrukturen, Datenformate und Gerätschaften sowie Bedarfe für Wissensbereiche im FDM-System erfasst worden. Basierend auf diesen Informationen sind ein Klassifikationssystem für die Komponenten der zu entwickelnden Technologie sowie ein verbundprojektspezifisches Vokabular abgeleitet, welches die Grundlagen für die Entwicklung von Ontologien darstellen. Die entwickelten Ontologien dienen dann in den FDM-Werkzeugen zur semantischen Vernetzung von Informationen und unterstützen bei der Implementierung von Wissensbereichen im FDM-System [5, 13]. Zur Erleichterung der Nachnutzung von Forschungsdaten nach Ende des Verbundprojektes werden Open-Source-Systeme eingesetzt. Das vorgeschlagene FDM Konzept sieht ein zentrales Datenmanagementsystem

(CKAN), ein Wissensmanagementsystem (Semantic MediaWiki, kurz: SMW) sowie eine Versionsverwaltung (Git) zur Versionierung von Software, Skripte und Konfigurationsdateien für Simulationen. Mittels eines projektspezifischen Vokabulars, eines Klassifikationssystems sowie Ontologien wird die semantische Vernetzung innerhalb der drei FDM-Säulen gewährleistet (vgl. Abbildung 1).

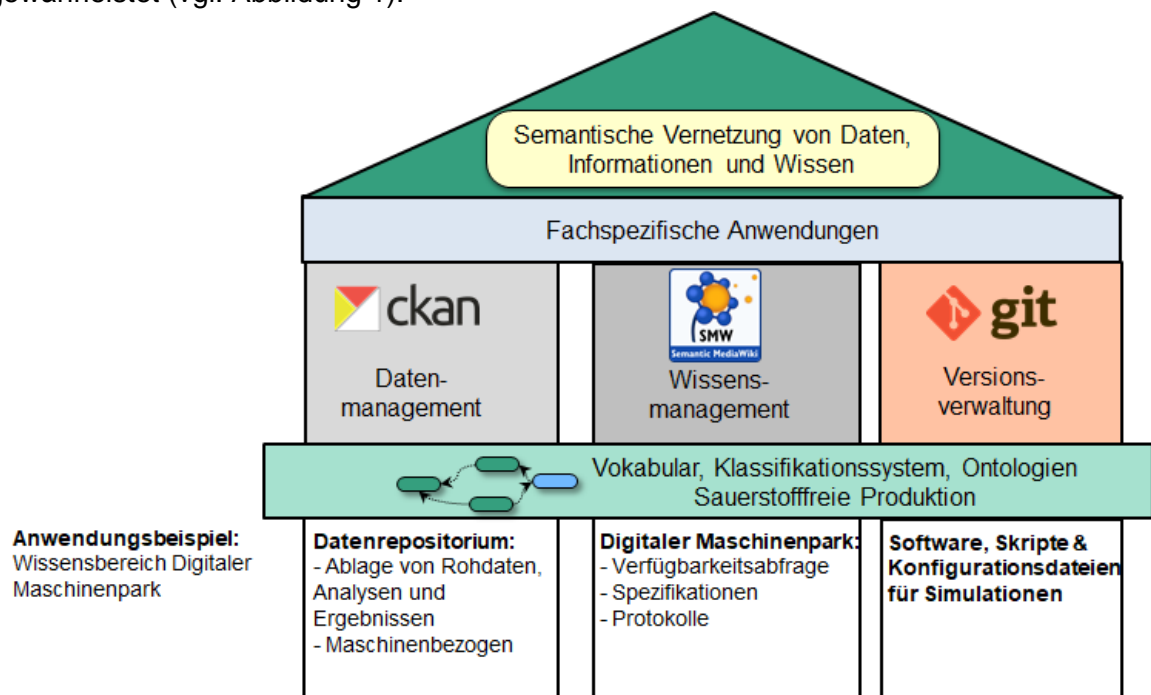


Bild 1: Bestandteile eines Forschungsdatenmanagementsystems

In der CKAN-Instanz werden Forschungsdaten gespeichert. In SMW wird das domänenspezifische Wissen zur Verfügung gestellt, z. B. in Form von Versuchs,- Test-, Simulationsprotokollen, bearbeiteten Werkstücken, verwendeten Werkstoffen oder Betriebsmitteln. Die zu annotierenden Inhalte werden in Wiki-Seiten organisiert und gespeichert. Die Struktur der Wiki-Seiten baut sich auf Basis einer domänenspezifischen Ontologie auf. In diesem Artikel wird ein bereits in SMW umgesetzter Wissensbereich, der digitale Maschinenpark, vorgestellt. Das nachfolgende Kapitel liefert eine umfangreiche Darstellung über die Entwicklung der semantischen Basis bzw. Ontologie für den Maschinenpark, deren Implementierung in SMW, sowie deren Verknüpfung mit den relevanten Datensätzen aus dem Datenmanagementsystem CKAN.

4. Anwendungsbeispiel - Digitaler Maschinenpark

Der digitale Maschinenpark soll die gesamte im Verbundprojekt verwendete technische Ausrüstung repräsentieren, einschließlich der Angaben zum Hersteller, der Modellbezeichnung, dem aktuellen Standort sowie der dafür verantwortlichen Person/Organisation. Jede Ausrüstung soll mit den relevanten Protokollen aus den Forschungsprozessen sowie den untersuchten Werkstücken semantisch verknüpft werden und in SMW als ein Wissensbereich den Forschenden zur Verfügung gestellt werden. In CKAN werden die Daten, welche mit der Ausrüstung erzeugt worden ist abgelegt und mit dem SMW verknüpft werden, sodass die Entstehung der Datensätze nachvollzogen werden kann. So können Forschende direkt einsehen, welche Daten und Ergebnisse mit welchen Betriebsmitteln erzeugt worden sind und können Redundanz in zu erforschenden Fragestellungen vermeiden. Das Wissensmanagementsystem ermöglicht darüber hinaus die semantische Abfrage von Wissen, wie beispielsweise nach bereits durchgeführten Untersuchungen an bestimmten Maschinen und verwendeten Materialien.

Zur Abbildung des digitalen Maschinenparks in SMW wird eine Ontologie entwickelt, das auf die Basisanforderungen der Forschenden beruht. Auf Grundlage der Basisanforderungen wurden der Begriffsrahmen der Ontologie sowie die Art der semantischen Beziehungen zwischen ihrer Entitäten identifiziert und folgende Kompetenzfragen definiert:

1. Welche Maschinen gibt es?
2. Welche Werkzeuge gibt es?
3. Welche Betriebsmittel sind unter der Verantwortung welcher Person und welcher Organisation?
4. Welches Betriebsmittel wird im Anwendungskontext *Test* verwendet?
5. Welches Betriebsmittel wird im Anwendungskontext *Experiment* verwendet?
6. Welches Betriebsmittel wird im Anwendungskontext *Handhabung und Transport* verwendet?
7. Welches Betriebsmittel wird im Anwendungskontext *Produktion* verwendet?
8. Welches Werkzeug wird mit der Maschine X verwendet?
9. Was ist der Standort des Betriebsmittels X, der unter Verantwortung des Projektes X steht?

4.1. MATO-Ontologie zur semantischen Beschreibung des digitalen Maschinenparks

Bei der Entwicklung einer semantischen Basis für den Maschinenpark wurden etablierte Ontologien berücksichtigt, die verschiedene Arten von Betriebsmitteln auf unterschiedlichen Abstraktionsebenen repräsentieren. Die Auswahl wurde auf Grundlage folgender Überlegungen getroffen: Die SOSA/SSN [14] enthält z. B. relevante Klassen-Entitäten (`ssn:System`, `sosa:Sensor` und `sosa:Actuator`), die jedoch von einem hohen Abstraktionsgrad sind. Überdies schließt SOSA/SSN das so genannte *System Capability Module* mit einer Reihe von System-Charakteristika ein. Diese Klassen und Properties eignen sich nicht für die projektspezifische Wiederverwendung, da der digitale Maschinenpark, zum einen, konkrete Betriebsmittelarten abbilden soll. Zum anderen, sollen sie sich über die Angaben zum Hersteller, Modell und Standort charakterisieren. SOSA/SSN kann daher nur für das Mapping mit anderen (Upper-)Ontologien, wie BFO [15], verwendet werden.

SAREF4INMA [16] enthält eine Super-Klasse für Produktionsbetriebsmittel und mehrere Klasse für diverse Betriebsmittelarten. Da der geplante digitale Maschinenpark weitere Anwendungsfelder als nur Produktion abdecken soll, kann SAREF4INGMA als Anregung für die Erfassung von Labels für individuelle Betriebsmittel dienen. MSDL [17] fokussiert auf Abbildung von Services im Produktionsbereich. Zwar liefert sie use-case-relevante Super- und Unterklassen (`tool` und `equipment`), diese unterscheiden sich in ihrer Semantik deutlich von den projektspezifischen Konzepten: Die Klasse `equipment` schließt `tool` aus, und der gegenseitige Bezug besteht nur auf der Unterklassen-Ebene. Im geplanten digitalen Maschinenpark Werkzeuge sollen Maschinen neben Werkzeugen die Unterarten von den Betriebsmitteln repräsentieren. MASON [18] beschränkt sich auf die Darstellung der Werkzeugmaschinen, die eine Relation zur Werkzeug-Klasse haben. Vergleichbare Darstellung liefert die *CDM-Core* [19], da sie die MASON-Klassen importiert. Dies weicht von den Anforderungen an den digitalen Maschinenpark, da darin -abhängig vom anzuwendenden Verfahren- unterschiedliche Werkzeugarten mit unterschiedlichem Maschinenarten sollen kombiniert werden können. Als Anregung für die Erfassung der Definitionen für Ontologie-Entitäten wurden folgende normative Dokumentation verwendet: *RICHTLINIE 2006/42/EG* [20] und *DIN EN 50370-2* [21]. Auf der Grundlage der Kompetenzfragen und unter der Berücksichtigung oben erwähnter Ontologien und Dokumenten, wurde die *Maschine and Tool Ontology* (MATO) entwickelt (Bild 2). Die MATO-Super-Klasse `mato:Equipment`

repräsentiert eine generische Entität, die für unterschiedliche oder für eine Kombination von unterschiedlichen Geräten, Maschinen, Anlagen und Werkzeugen steht und die erforderlich ist, um den beabsichtigten Zweck zu erreichen. Da der reale Maschinenpark zwei zentrale Betriebsmittelarten umfasst, wurde `mato:Equipment` um zwei Unter-Klassen erweitert: Die `mato:Machine` repräsentiert ein Betriebsmittel in Form von Geräten, Maschinen, Anlagen oder deren Kombination, die entweder autonom ist oder aus Teilen und Komponenten besteht, die als integrales Ganzes funktioniert und den beabsichtigten Zweck selbst erfüllt. Die `mato:Tool` steht für ein Betriebsmittel, welches ausschließlich dazu bestimmt ist, in ein anderes Betriebsmittel eingebaut zu werden, um den bestimmungsgemäßen Zweck zu erfüllen und welches die Funktion dieses Betriebsmittels weder verändert noch erweitert. Zwischen `mato:Machine` und `mato:Tool` besteht eine semantische Relation `mato:usesTool`, die es ermöglicht, einen Werkzeug, der in einem bestimmten Prozess in Kombination mit einer bestimmten Maschine verwendet werden soll, dieser Maschine semantisch zuzuordnen.

MATO-Betriebsmittel werden mittels folgender Attribute charakterisiert: den Herstellernamen (`mato:hasManufacturer`), die Modellbezeichnung (`mato:hasModel`), die Abbildung (`foaf:depiction`), eine kurze Beschreibung (`mato:hasDescription`) und den Standort (`mato:atLocation`). Überdies kann durch die Property `mato:isOfMaterialEntityType` ausgedrückt werden, ob das Betriebsmittel vom Type `mato:Hardware` oder `mato:Software` ist. Eine Relation zwischen Betriebsmittel und der dafür zuständigen Person und Organisation (`mato:ResponsibleForEquipment`), der Property `mato:underResponsibilityOf`. Die Property `mato:employsEquipment` verknüpft semantisch die Klassen `mato:Equipment`, `mato:Machine` und `mato:Tool` mit deren Anwendungskontext (`mato:ApplicationContext`). Im verbundspezifischen Kontext sind vier Anwendungsarten gegeben, die in MATO durch folgende Individuen dargestellt sind: (1) `mato:Production`, (2) `mato:Handling_and_transportation`, (3) `mato:Inspection` und (4) `mato:Test`.

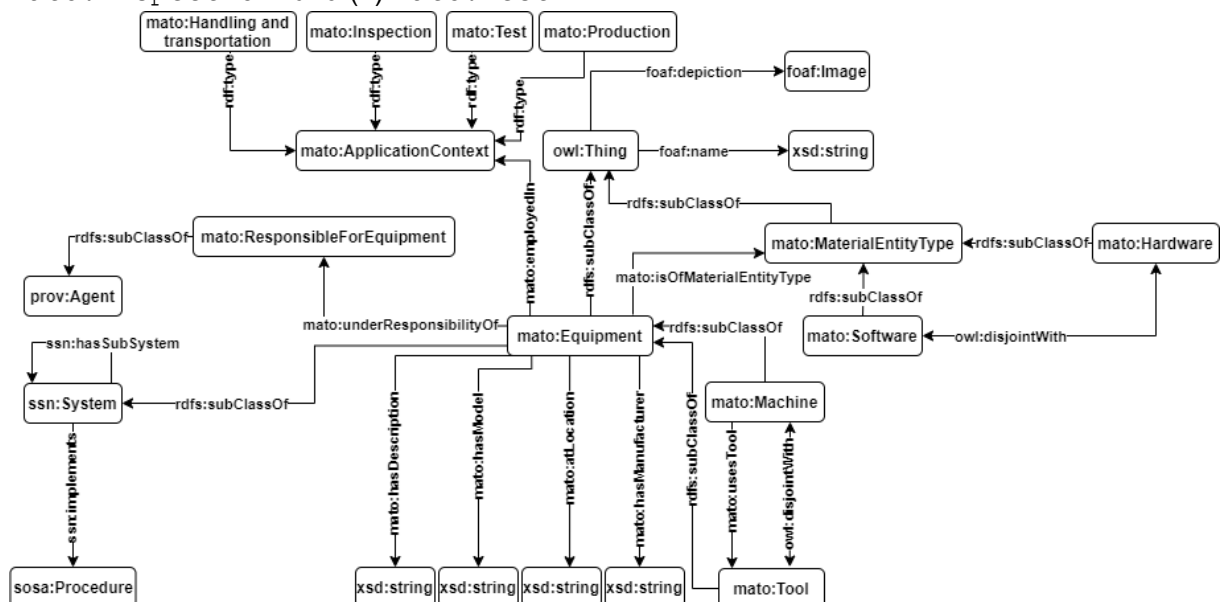


Bild 2: Grafische Abbildung der *Machine and Tool Ontology* (MATO)

Die Entwicklung der MATO erfolgte unter Berücksichtigung der FAIR-Prinzipien. So wurde *Reusability*, *Interoperability*, *Accessibility* und *Findability* durch die Wiederverwendung bereits existierender Ontologien und durch die Speicherung der MATO im maschineninterpretierbaren TTL-Format in dem öffentlich zugänglichen Git-Repository TIB Onto sichergestellt. MATO ist unter folgendem Link verfügbar: <https://github.com/tibonto/mato/blob/main/mato.ttl>.

4.2. Realisierung des digitalen Maschinenparks in Semantic MediaWiki

Auf Basis der Struktur der MATO ist es möglich, die maschinen-, und werkzeugrelevanten Informationen zu sammeln, zu strukturieren, semantisch zu annotieren und daraus einen digitalen Maschinenpark aufzubauen. Die semantische Annotation von Betriebsmitteln erfolgt unter Verwendung von MATO-Klassen, -Individuen und -Properties. Dabei sind die Klassen `mato:Machine` und `mato:Tool` als SMW-Kategorien dargestellt. Instanzen dieser Klassen werden durch die einzelnen SMW-Seiten repräsentiert. Die den MATO-Klassen zugeordnete Properties aus der Ontologie wird zur Repräsentation von Eigenschaften der Betriebsmittel-Individuen aus SMW-Seiten und zur Suche nach diesen verwendet.

4.3. CKAN-Erweiterung zur Kommunikation mit Semantic MediaWiki

Zur Kontextualisierung von Forschungsdaten, die in CKAN hochgeladen werden, ist es nötig Datensätze mit Metadaten zu beschreiben und die Suche bzw. Abfrage nach diesen zu ermöglichen. In CKAN können den Datensätzen Metadaten wie Projektverantwortlichkeit, Organisationszugehörigkeit, Lizenzierung etc. abgebildet werden. Indem den Datensätzen Betriebsmittel, die im Wissensmanagementsystem SMW eigene Wissensbereiche repräsentieren und Protokolle über bereits durchgeführte Experimente enthalten, zugeordnet werden, gelingt es, den Forschenden zu ermöglichen die Entstehung der Daten nachzuvollziehen. Bild 3 veranschaulicht den Ablauf für die Verknüpfung von Betriebsmittel, die in SMW gespeichert sind mit Datenressourcen in CKAN. Zur Realisierung der Verknüpfung ist es notwendig gewesen CKAN mit einer zusätzlichen Funktionalität bzw. CKAN-Erweiterung und einem entsprechendem Plugin, hier als *media_wiki* bezeichnet auszustatten. Das *media_wiki-Plugin* ermöglicht es Forschern, die Geräte auf SMW mit einer Datenressource in CKAN zu verknüpfen. Zunächst erstellt ein Forscher einen Datensatz auf CKAN, indem er die Metadaten des Datensatzes eingibt und die zugehörigen Daten hoch lädt. Nach dem Speichern des Datensatzes führt CKAN das *media_wiki-Plugin* aus.

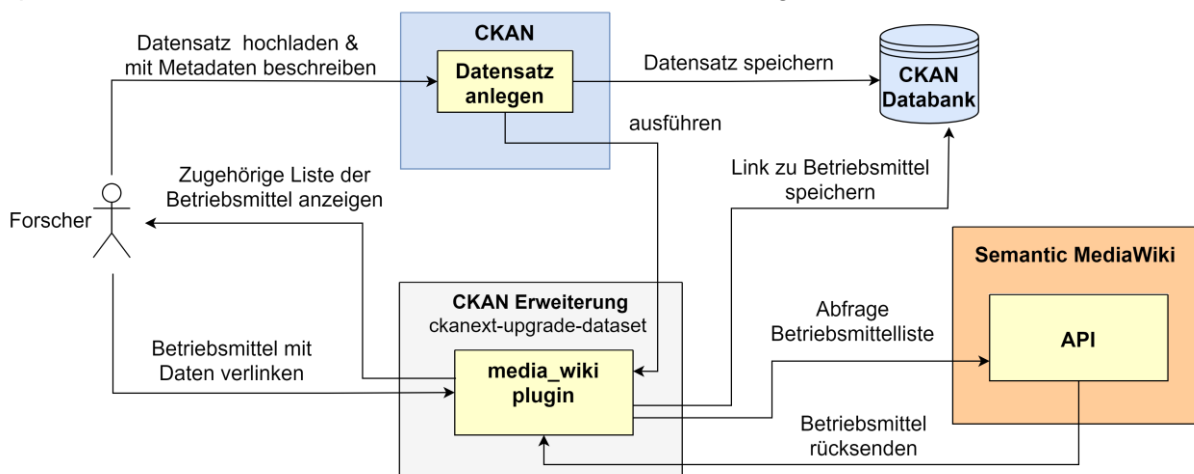


Bild 3: Verknüpfung von Betriebsmittel in Semantic MediaWiki mit Datenressourcen in CKAN

Das *media_wiki-Plugin* sendet eine Abfrage an die SMW API und ruft die Liste der Geräte ab und zeigt sie dem Forscher an. Der Forscher wählt das Zielgerät für jede Datei im hochgeladenen Datensatz aus. Schließlich speichert das *media_wiki-Plugin* die URL des ausgewählten Geräts in der CKAN-Datenbank. Für andere Forschende wird die URL zu den Geräten im SMW auf der entsprechenden Seite des Datensatzes in CKAN abgebildet. Auf diese Art und Weise gelingt es Datensätzen Betriebsmittel, die in SMW abgelegt sind, zuzuordnen. So können Forschende den Entstehungskontext der Daten nachvollziehen.

4.4. Beispiel Digitaler Maschinenpark - Kommunikation zwischen CKAN und SMW

Durch die in den Abschnitten 4.1, 4.2 sowie 4.3 beschriebenen Erweiterungen der Open-Source Systeme CKAN und SMW ist ein an einen digitalen Maschinenpark gekoppeltes Datenmanagementsystem realisiert worden. In Bild 4 sind die jeweiligen User-Interfaces zu sehen. Auf der linken Seite ist ein Ausschnitt aus CKAN abgebildet. Nach erfolgreicher Ablage von Daten (hier: Ressourcen) in CKAN kann der Nutzer eine Maschine aus dem zuvor in SMW erstellten Maschinenpark auswählen und so den Datensatz weiter kontextualisieren. Auf der rechten Seite ist ein Ausschnitt aus der individuellen Maschinenseite in SMW abgebildet. Wie bereits in Abschnitt 4.2 beschrieben, kann der Nutzer eine semantische Abfrage in SMW nach den Eigenschaften der Maschine durchführen.

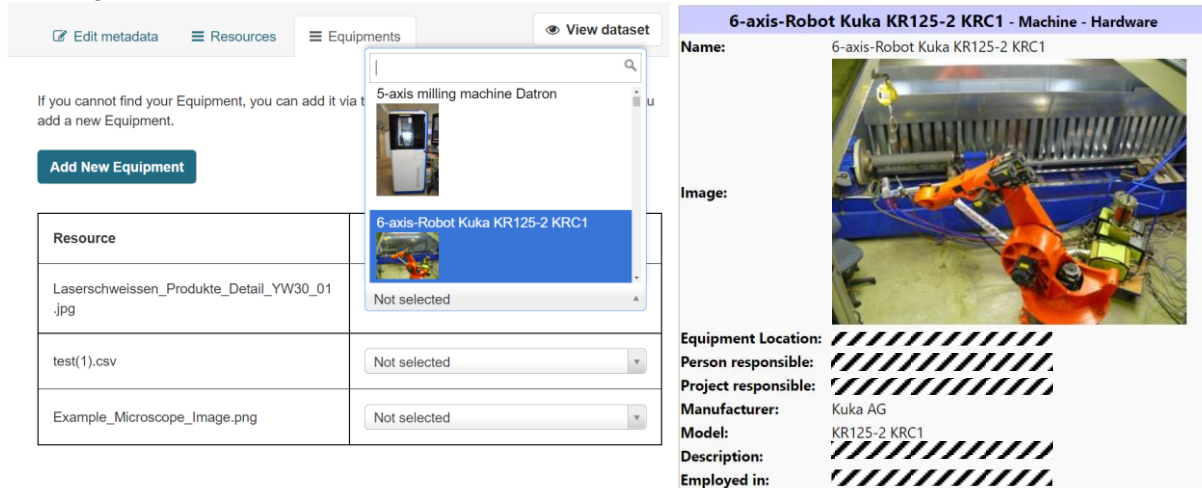


Bild 4: Links - Ausschnitt CKAN (Verlinkung eines Datensatzes mit dem Maschinenpark); Rechts - Ausschnitt individuelle Maschinenseite in SMW

5. Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit sind nach einer grundlegenden Motivation der Entwicklung und Nutzung von FDM-Systemen im Rahmen von Verbundprojekten, unterschiedliche bereits vorhandene Systeme zur Realisierung eines FDM vorgestellt worden. Die Nutzung von FDM-Systemen sind insbesondere in großen interdisziplinären Forschungsprojekten wesentlich, da hier große Mengen heterogener experimenteller Daten entstehen und deren kompetente und umfängliche Verarbeitung und Analyse zur Durchführung einer umfassenden Forschung erforderlich ist. Der Beitrag bietet ein Konzept zur Realisierung eines FDM-Systems nach den FAIR-Daten-Prinzipien mit Hilfe von Open-Source-Systemen und deren Anpassung an projektspezifische Anforderungen mit Hilfe von Klassifikationssystemen, kontrolliertem Vokabular sowie Ontologien. Anhand des Beispiels eines digitalen Maschinenparks im Rahmen des SFB1368 ist die Umsetzung eines Wissensbereiches im vorgeschlagenen FDM-System vorgestellt worden. Die vorgestellten und diskutierten Systeme können auch in anderen Verbundprojekten zum Einsatz kommen. Die Realisierung weiterer Wissensbereiche sowie Bereitstellung eines methodischen Vorgehensmodells zur Umsetzung derartiger FDM-Systeme wird Bestandteil zukünftiger Beiträge sein.

Danksagung

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 394563137 – SFB 1368.

Literaturverzeichnis

- [1] Kapogiannis, Georgios; Sherratt, Fred: Impact of integrated collaborative technologies to form a collaborative culture in construction projects. In: Built Environment Project and Asset Management 8 (2018), Nr. 1, S. 24-38.
- [2] Sandfeld, Stefan et al.: Strategiepapier Digitale Transformation in der Materialwissenschaft und Werkstofftechnik. Frankfurt: Deutsche Gesellschaft für Materialkunde e.V., 2018.
- [3] Amorim, Ricardo et al.: A comparison of research data management platforms: architecture, flexible metadata and interoperability. In: Universal Access in the Information Society, 16 (2017), Nr. 4, S. 851-862.
- [4] Wilkinson, Mark D. et al.: The FAIR Guiding Principles for scientific data management and stewardship. In: Scientific Data 3 (2016), Nr. 160018.
- [5] Mozgova, Iryna et al.: Research Data Management System for a large Collaborative Project. In: DS101: Proceedings of NordDesign 2020, Lyngby: The Design Society, 2020.
- [6] Koppe, R. et al.: O2A: A Generic Framework for Enabling the Flow of Sensor Observations to Archives and Publication. In: OCEANS 2015, Genova, 2015, S. 1-6.
- [7] Norm DIN 4000-1, 2019. Sachmerkmal-Listen – Begriffe und Grundsätze.
- [8] Pellegrini, Tassilo; Sack, Harald; Auer, Sören (Hrsg.): Linked Enterprise Data - Management und Bewirtschaftung vernetzter Unternehmensdaten mit Semantic Web Technologien. Berlin Heidelberg: Springer-Verlag, 2014.
- [9] Vandenbussche, Pierre-Yves et al.: Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web. In: Semantic Web 8 (2014), Nr. 3, S. 437-452.
- [10] Maier, Hans Jürgen et al.: Towards Dry Machining of Titanium-Based Alloys: A New Approach Using an Oxygen-Free Environment. In: Metals 10 (2020), Nr. 9, 1161.
- [11] Szafarska, Maik; Gustus, Rene; Maus-Friedrichs, Wolfgang: Sauerstofffreier Transport, Präparation und Transfer von Materialproben für die Oberflächenanalytik. In: Clausthaler Zentrum für Materialtechnik (Hrsg.): Tagungsband 4 . Symposium Materialtechnik. Düren: Shaker Verlag, 2021, S. 829–839.
- [12] Effertz Eva: The Funder's Perspective: Data Management in Coordinated Programmes of the German Research Foundation (DFG). In: Curdt, Constanze; Bareth, Georg (Hrsg.): Proceedings of the Data Management Workshop 29.-30.10.2009 Cologne, Köln: Geographisches Institut der Universität zu Köln, 2010, S. 35-38.
- [13] Sheveleva, Tatyana et al.: Development of a Domain-Specific Ontology to Support Research Data Management for the Tailored Forming Technology. Procedia Manufacturing 52 (2020), Nr. 1, S. 107-112.
- [14] Armin Haller et al.: Semantic Sensor Network Ontology. W3C Editor's Draft 23 July 2021. [online]: <https://w3c.github.io/sdw/ssn/>, zugegriffen am 05.07.2021.
- [15] Arp, Robert; Smith, Barry; Spear, Andrew: Building Ontologies With Basic Formal Ontology. Cambridge: MIT Press, 2015.
- [16] Daniele, Laura: SAREF4INMA: an extension of SAREF for the industry and manufacturing domain, 2019 [online]: <https://saref.etsi.org/saref4inma/v1.1.2/>, zugegriffen am 30.07.2021.
- [17] Ameri, Farhad; Dutta, Debasish: An Upper Ontology for Manufacturing Service Description, In: Proceedings of the ASME 2006 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference 3 (2006), S. 651-661.
- [18] Lemaignan, Séverin et al.: MASON: A proposal for an ontology of manufacturing domain. In: IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06), 2006, S. 195-200.
- [19] Mazzola, Luca et al.: CDM-Core: A Manufacturing Domain Ontology in OWL2 for Production and Maintenance. In: IC3K 2016: Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management 2016, S. 136–143.
- [20] Richtlinie 2006/42/EG, 2006. Über Maschinen und zur Änderung der Richtlinie 95/16/EG
- [21] Norm DIN EN 50370-2, 2003. Elektromagnetische Verträglichkeit (EMV) - Produktfamilienorm für Werkzeugmaschinen - Störfestigkeit.