

# Semantic Representation of Low-Cycle-Fatigue Testing Data Using a Fatigue Test Ontology and ckan.kupferdigital Data Management System

Hossein Beygi Nasrabadi,\* Thomas Hanke, and Birgit Skrotzki

Addressing a strategy for publishing open and digital research data, this article presents the approach for streamlining and automating the process of storage and conversion of research data to those of semantically queryable data on the web. As the use case for demonstrating and evaluating the digitalization process, the primary datasets from low-cycle-fatigue testing of several copper alloys are prepared. The fatigue test ontology (FTO) and ckan.kupferdigital data management system are developed as two main prerequisites of the data digitalization process. FTO has been modeled according to the content of the fatigue testing standard and by reusing the basic formal ontology, industrial ontology foundry core ontology, and material science and engineering ontology. The ckan.kupferdigital data management system is also constructed in such a way that enables the users to prepare the protocols for mapping the datasets into the knowledge graph and automatically convert all the primary datasets to those machine-readable data which are represented by the web ontology language. The retrievability of the converted digital data is also evaluated by querying the example competency questions, confirming that ckan.kupferdigital enables publishing open data that can be highly reused in the semantic web.

facilitates advanced capabilities in semantic query, exchange, retrieval, integration, visualization, and analysis of data across diverse networks.<sup>[2]</sup> This approach not only fosters data-driven research methodologies, such as machine learning and deep learning,<sup>[3]</sup> but also yields significant industry benefits for product development, discovering new materials, enhancing production efficiency, and even reducing manufacturing costs.<sup>[4]</sup>

As a use case for research data in materials science, the digitalization of low-cycle-fatigue (LCF) datasets is being evaluated in this study. LCF is a material testing method that assesses the durability of materials when they are subjected to repeated loading and unloading cycles at relatively high-stress levels.<sup>[5]</sup> In materials engineering, it is essential to ensure the endurance and structural integrity of parts under cyclic loads, such as mechanical parts or aircraft components, as this prevents unexpected failure and confirms reliability and safety

in a variety of applications.<sup>[6]</sup>


Basically, two prerequisites are required for the digitalization of materials research data and converting them to FAIR machine-readable and semantically queryable datasets. 1) Ontologies: Ontology serves as a powerful tool for creating a coherent and meaningful knowledge structure by formulating a formal standardized knowledge representation.<sup>[7]</sup> It establishes a structured vocabulary and defines relationships between entities of various domains through classes, properties, constraints, individuals, and axioms.<sup>[7b,8]</sup> Web ontology language (OWL) is a powerful language for defining ontologies, and it often utilizes resource description frameworks (RDFs) to represent data through subject–predicate–object triples.<sup>[9]</sup> Until 2020, a few ontologies were developed in the materials testing domain (like creep,<sup>[10]</sup> tensile,<sup>[11]</sup> or fatigue<sup>[12]</sup> ontologies) that remained unused due to their incompatibility with testing datasets, lack of instances and process knowledge bases, and low level of details and relations for the description of the test process. By reusing the well-defined upper-level ontologies, some industrially applicable ontologies were introduced recently in the materials testing domain (like Brinell hardness,<sup>[13]</sup> Vickers hardness,<sup>[14]</sup> tensile test,<sup>[2,15]</sup> and mechanical testing ontology<sup>[2]</sup>) that provided higher interoperability with other existing ontologies and enabled the efficient description of mechanical testing datasets. Concerning

## 1. Introduction

In the era of digitalization and Industry 4.0, the publication of open and findable, accessible, interoperable, and reusable (FAIR) data, particularly in the materials science domain, marks a transformative shift.<sup>[1]</sup> Leveraging semantic web technologies

H. B. Nasrabadi, B. Skrotzki  
Department 5.2 - Metallic High-Temperature Materials  
Bundesanstalt für Materialforschung und -prüfung (BAM)  
Unter den Eichen 87, 12205 Berlin, Germany  
E-mail: hossein.beygi-nasrabadi@bam.de

T. Hanke  
Assessment of Materials and Lifetime Concepts  
Fraunhofer-Institut für Werkstoffmechanik (IWM)  
Wöhlerstraße 11, 79108 Freiburg im Breisgau, Germany

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adem.202400675>.

© 2024 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adem.202400675

fatigue testing, however, the previously reported examples<sup>[12]</sup> just provided the simple fatigue terminology classifications and there are not any standardized ontologies that allow efficient digitalization of fatigue testing datasets. 2) Data management systems for data mapping, conversion, and storage: The ontological representation of testing datasets is performed by linking the content of the datasets with their corresponding ontology entities (data mapping) and then converting them to RDF files.<sup>[2]</sup> Such files are the machine-readable version of the testing datasets which can be efficiently reused and queried in the web.<sup>[2]</sup> The data mapping and conversion can be done by writing scripts utilizing the RDF libraries of Java (Apache Jena), C# (dotNetRDF), and Python (RDFLib),<sup>[16]</sup> through the ontology editors (like Protégé, and TopBraid composer<sup>[17]</sup>), or using some tools like Mat-O-Lab toolchain,<sup>[2]</sup> Ontopanel,<sup>[13]</sup> OpenRefine,<sup>[18]</sup> CSV2RDF,<sup>[19]</sup> and ontop.<sup>[20]</sup> While script writing needs the programming skills of the user and is a time-consuming process for processing different types of datasets, ontology editors and tools provide graphical user interfaces and simplify the data mapping and RDF conversion processes. For example, the Mat-O-Lab toolchain provides a professional and powerful environment for the parsing, mapping, and RDF conversion of primary CSV and XLSX test data.<sup>[2]</sup> The Ontopanel tool can also graphically map the complex test datasheets in the process graphs and convert them to RDF files.<sup>[13]</sup> Although they can handle some aspects of the data extraction, mapping, transferring, and validation procedures, several restrictions are associated with the tools introduced so far. Limitations with the processing of heterogeneous primary datasets (e.g., various file types, structures, languages, and units), confined ontology libraries or impossibility of importing desired ontologies, as well as lack of efficient data management systems for storing the converted files can be named as the main restrictions of the current tools in the processes of data conversion and digitalization.

This research aims to address the RDF conversion and storage of semantically queryable data on the web by developing the requirements for the digitalizing process of the use case LCF datasets. The detailed fatigue knowledge graph and ontology are modeled according to the testing standard and using the appropriate top-level ontologies. The developed ontology not only should be highly compatible with testing datasets and other existing ontologies but also should be efficiently applicable in the digital environment. The ckan.kupferdigital data management system is also developed in such a way that streamlines and automates the processes for data mapping, RDF conversion, and data storage. To evaluate how the developed ontology and data management systems can lead the efficient data mapping and conversion, the retrievability of the converted data is also evaluated by querying the following exemplary competency questions. 1) Which materials were tested by the LCF process? 2) Under which LCF conditions (e.g., strain range and ratio, strain rate, cycle time, temperature) were the materials tested? 3) What was the measured fatigue life of the materials?

The findings of this study show that ckan.kupferdigital can efficiently link primary datasets and fatigue testing ontology (FTO) then convert them to RDF data that are fully queryable in the semantic web. The ckan.kupferdigital offers an automated and user-friendly approach to both industry and academia for efficiently making their research data FAIR. Here, the converted

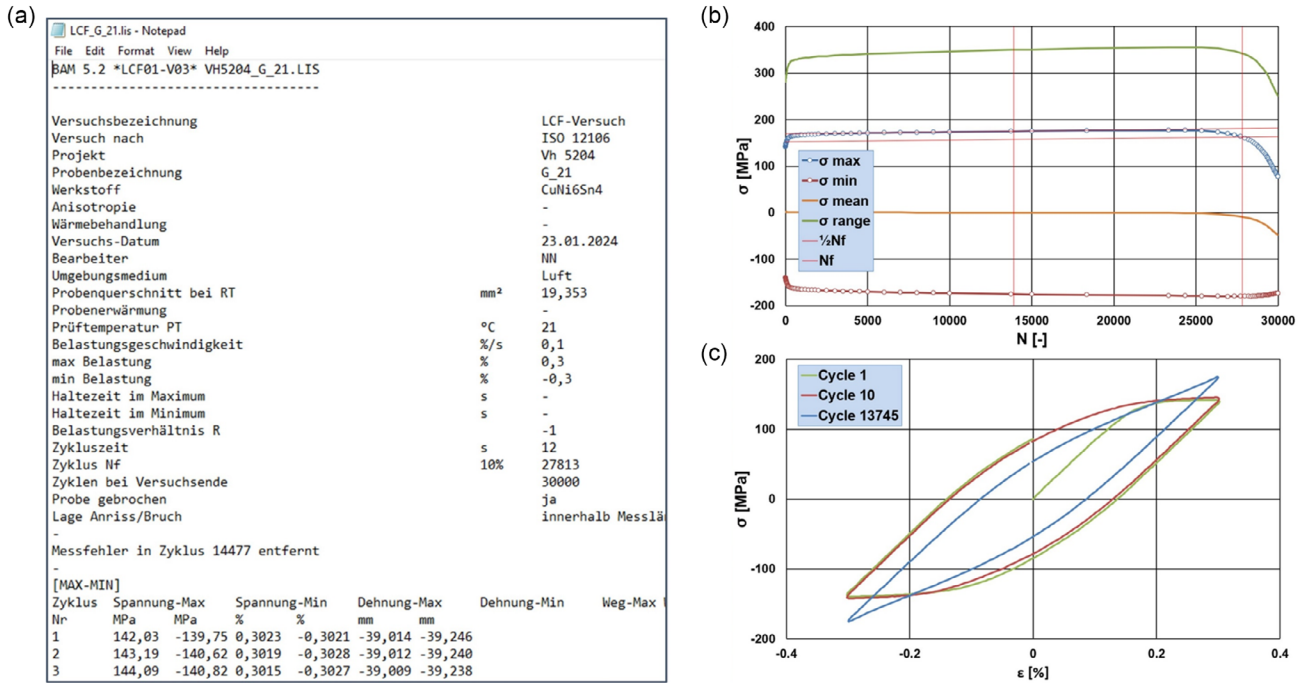
datasets are the comprehensive machine-readable files that are available in open-access data repositories and can be easily queried and reused by the data users. Collecting such FAIR data from multiple sources allows both industries and academia to produce valuable datasets for various data-driven and product development projects.

## 2. LCF Testing Datasets of Copper Alloys

To prepare the LCF datasets, several copper alloys have been tested by the standard strain-controlled fatigue test procedure<sup>[5b]</sup> and at different strain ranges. The results of these measurements were reported by the primary and secondary LCF datasets.<sup>[21]</sup> The primary datasets are the raw data from the test, which are provided directly by the material testing laboratory. These data were represented in German language and within the CSV-type files (Figure 1a). In a follow-up study, it will be demonstrated how data digitization converts such use case data into machine-readable data of standard structure and language that can be efficiently queried on the web. The primary LCF test data in this study consisted of two parts. The upper part represents various metadata about the LCF test (used standard, project ID, test date, operator), test piece characteristics (ID, chemical composition, cross-sectional area), testing parameters (atmosphere, temperature, strain rate, strain range, strain ratio, cycle time), and measured test values (like fatigue life). The bottom part of the primary dataset reports the variation of stress, strain, and displacement during the LCF. These tabular datasets are used for plotting the variation of different LCF variables. For example, Figure 1b illustrates that during the LCF test, the maximum stress (blue curve) in the sample increased rapidly during the first few cycles and then slightly but continuously. After exceeding 25 000 cycles, a rapid decline of the max. stress is observed resulting in the sample's failure after 30 000 cycles. In this regard, the fatigue life ( $N_{f10\%} = 27\,813$ ) was determined using the failure criterium of 10% load drop of the maximum stress concerning the slope of the stabilized deformation region (thin red line), and  $1/2N_f$  ( $\approx 13\,745$ ) represents half of the fatigue life.

Indeed, the cyclic variation of stress and strain can be explained by the hysteresis loops of Figure 1c. During each loading–unloading cycle, the material experiences elastic deformation followed by plastic deformation, resulting in a hysteresis loop. Here, the LCF tests were conducted in the strain-controlled mode which creates constant deformation amplitude at each cycle (e.g., from  $-0.3\%$  to  $0.3\%$  for sample G\_21). As shown in Figure 1c, the stress increases proportionally in the first fatigue cycles adhering to Hooke's law. The accumulation of plastic deformation at higher cycles leads to nonlinear stress–strain behavior and subjects the materials (in the shown example) to strain hardening, causing increased resistance to deformation. Ultimately, stress concentration at localized points may result in crack initiation and propagation, leading to failure.

The LCF secondary dataset has been consequently created by collecting the main measurement metadata and analyzing data from all LCF tests in an XLSX file.<sup>[21]</sup> As the example column headings of the LCF secondary dataset can name the identifier, composition, length, and original cross-section area of the test piece, testing machine identifier, test temperature, strain rate,



**Figure 1.** a) Example of LCF primary dataset for the test piece with ID G\_21 and composition CuNi6Sn4 (wt%), b) variation of stress response during the LCF testing of sample G\_21 ( $\sigma_{max}$  and  $\sigma_{min}$  represent the extremum stress values within each stress–strain hysteresis loop,  $\sigma_{mean} = \sigma_{max} + \sigma_{min}/2$ , and  $\sigma_{range} = \sigma_{max} - \sigma_{min}$ ), and c) stress–strain hysteresis loops of cycles #1, #10, and #13 745 during the LCF testing of sample G\_21.

maximum/minimum strain, strain ratio, cycle time, fatigue life, fracture cycles, and fracture location. **Table 1** represents a small part of the LCF secondary dataset. This table is used to plot fatigue life variation by strain amplitude, estimate fatigue strength and service life of materials, and design engineering components with efficient fatigue life. For example, **Figure 2** shows that when comparing CuSn8Ni2, CuZn23Si2.5, and CuNi6Sn4 alloys at a constant strain amplitude, CuSn8Ni2 has the highest fatigue life.

### 3. Fatigue Test Ontology (FTO)

FTO has been modeled to provide a standardized structure for representing, organizing, and defining the fatigue testing knowledge, hence leading to its dataset interoperability in digital systems. The development of FTO has been performed by graphical designing the fatigue testing knowledge graph in the diagrams.net<sup>[22]</sup> tool and then converting its resulting XML schema into the OWL formats like RDF or turtle syntax (TTL) using the Ontopanel plugin.<sup>[13]</sup> Here, the fatigue testing knowledge graph should semantically represent the testing process and its relationships to the materials, equipment, process history, and data analysis characteristics. To address these issues, **Figure 3–5** respectively show the different parts of FTO which modeled the entities of the test piece, equipment and measurement procedure, and analysis of testing data.

The graphical models from different parts of FTO were developed in the distinct T- and A-boxes. The T-Box represents the

**Table 1.** Part of the LCF secondary dataset for various copper alloys (the full dataset can be found in).<sup>[21]</sup>

Sample	$e_{max}$ [%]	$e_{min}$ [%]	Fatigue life, $N_{f10\%}$ (–)
G_21	0.30	–0.30	27 813
G_22	0.45	–0.45	8183
G_23	0.40	–0.40	7868
G_24	0.35	–0.35	12 422
G_25	0.50	–0.50	3291
G_26	0.45	–0.45	5617
G_27	0.40	–0.40	8788
H_21	0.55	–0.55	21 040
H_22	0.45	–0.45	35 640
H_23	0.60	–0.60	13 463
H_24	0.35	–0.35	run out <sup>a)</sup>
H_25	0.50	–0.50	29 038
H_26	0.45	–0.45	65 372
H_27	0.40	–0.40	run out <sup>a)</sup>
L_21	0.50	–0.50	9726
L_22	0.45	–0.45	15 576
L_24	0.35	–0.35	50 749
L_25	0.40	–0.40	20 789

<sup>a)</sup>More than 100 000 cycles were defined as “run out”.

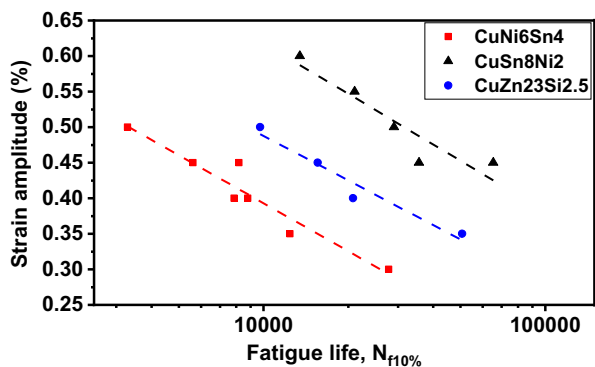


Figure 2. Fatigue life variation by strain amplitude from LCF testing of different copper alloys.

ontology classes, which mainly come from basic formal ontology (BFO),<sup>[23]</sup> industrial ontology foundry (IOF),<sup>[24]</sup> and material science and engineering ontology (MSEO)<sup>[25]</sup> upper-level ontologies. Addressing ISO/IEC 21838-2:2021 standard,<sup>[26]</sup> BFO is utilized as the top-level ontology for modeling FTO due to its foundational principles and simplicity. BFO not only provides a well-defined and minimalistic framework that facilitates semantic interoperability, consistency, and ontological clarity but also emphasizes capturing fundamental categories and relationships.<sup>[2]</sup> The ontologies which are developed based on BFO can be easily merged with a variety of ontologies making them highly interoperable. Furthermore, we utilized IOF core ontology since it is highly compatible with BFO and offers a specialized framework that aligns with industry-specific concepts, providing lots of manufacturing and industrialization concepts

for modeling the FTO. MSEO was also particularly utilized for expressing specific material-related quantities and units. Eventually, all the newly created ontology entities were expressed with the FTO namespace.<sup>[27]</sup>

The A-Box of FTO modeled its knowledge graph according to the fatigue testing procedure described by ISO 12106:2017 standard.<sup>[5b]</sup> Here, the testing standard determined the scientific and technical logic for designing the knowledge graph, as well as the level of detail for the process description. Furthermore, the test standard is the most important resource for preparing the ontology vocabulary, since it provides the specific names and definitions for the different concepts of the processes and orients the structuring and categorization of test process entities based on the information that has already been agreed by an expert standardization committee.<sup>[8]</sup> In addition, developing the ontologies based on the materials testing standards ensures their acceptance and applicability in the industry.<sup>[2]</sup>

The graphical design of knowledge graphs was performed utilizing the Ontopanel library,<sup>[13]</sup> where each shape has a particular definition. Here, all white rectangles of the A-Box are *individuals* that are connected by the brown dashed lines (*types*) to the colored rectangles of the T-Box (*classes*). The arrows between white rectangles are *object properties* that express the semantic relationships between different individuals. Furthermore, any data related to the testing process are expressed within the purple rectangles (*data values*) and connected to their related individual by the *data properties*, arrows between white and gray rectangles. Eventually, the gray rectangles of A-Box are individuals that are planned for mapping the entities of testing primary/secondary datasets. All these individuals are a type for one of the *cco: InformationContentEntity* classes. The above-mentioned shapes and their embedded texts are subsequently converted to TTL

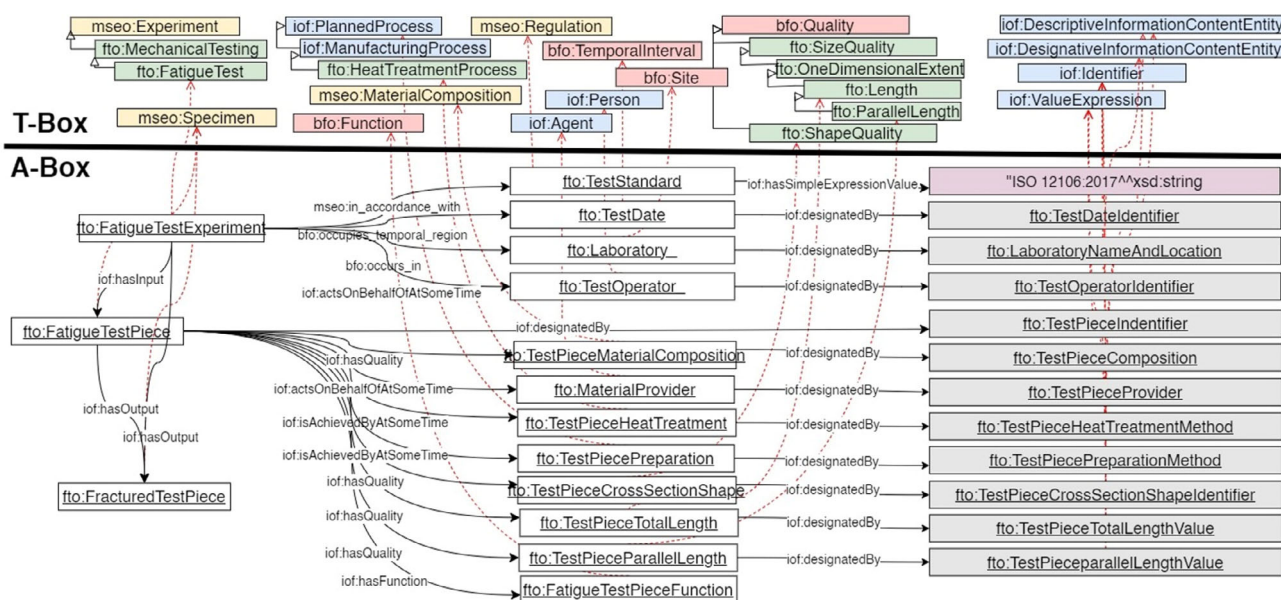
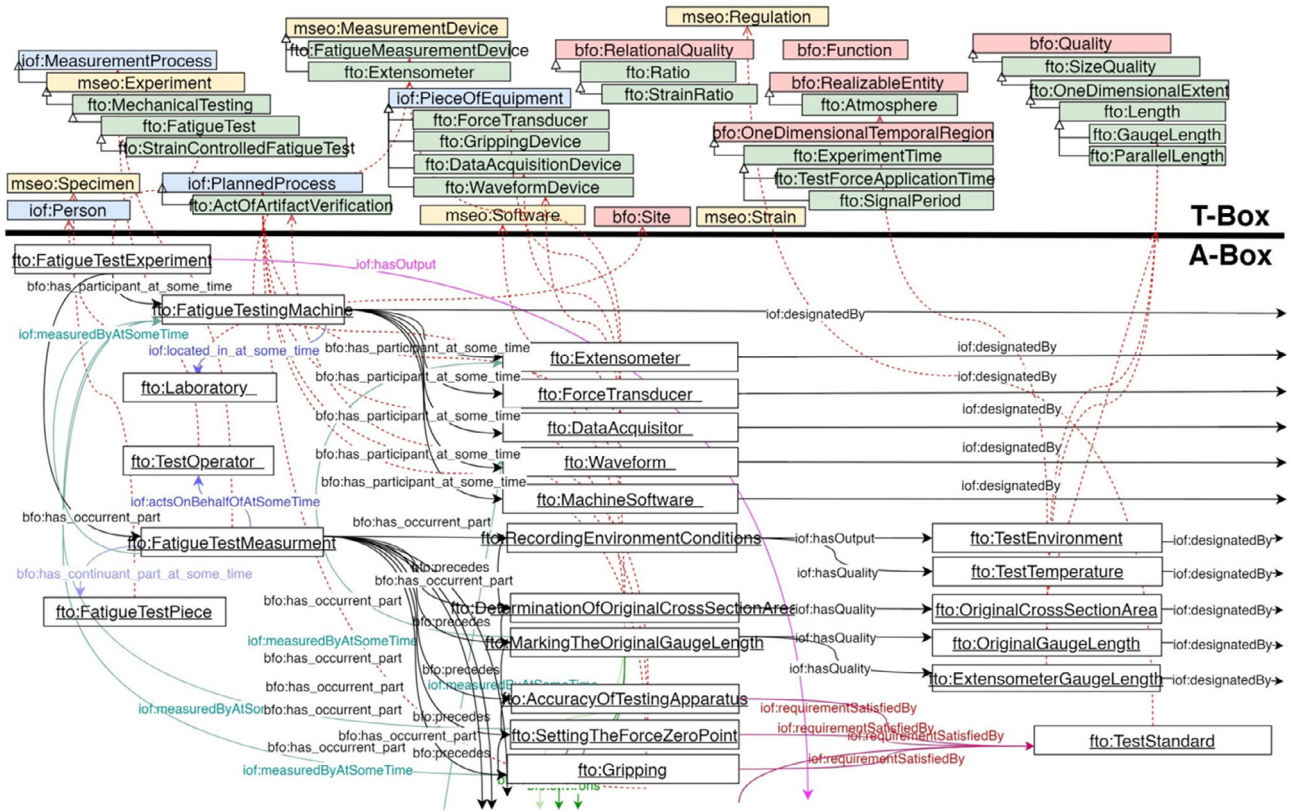
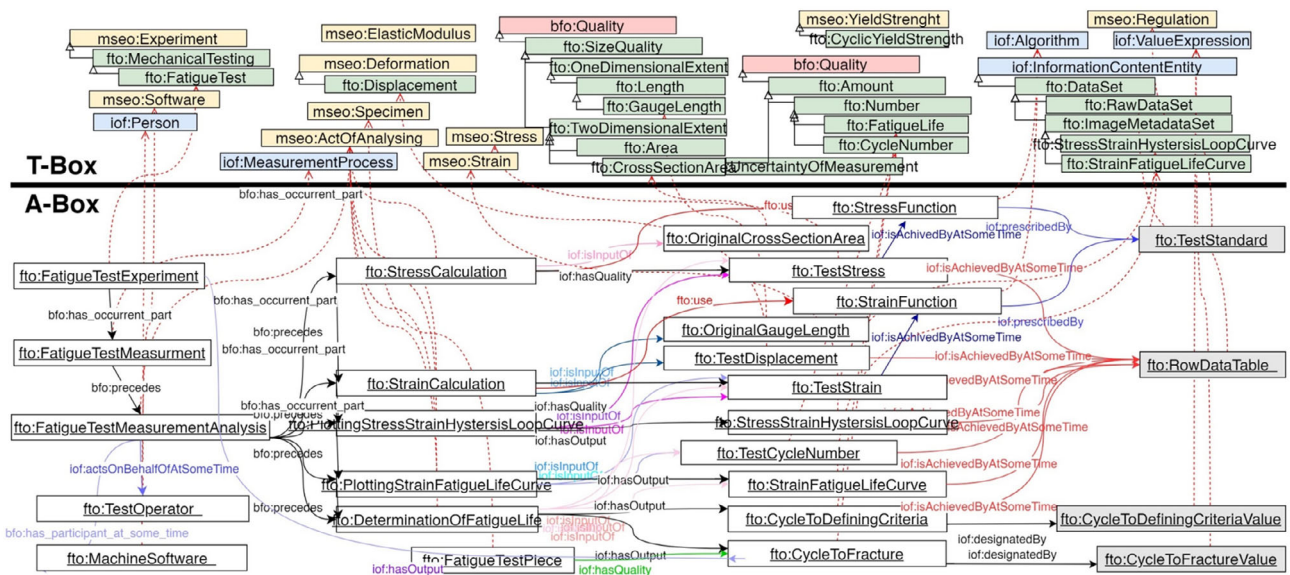


Figure 3. The graphical schema from the “provenance” and “test piece” parts of FTO. The T-Box shows the classes that were mainly driven from BFO, IOF, or MSEO upper-level ontologies (red-, blue-, and yellow-colored rectangles, respectively), while the new classes from FTO were distinguished by green color. The A-Box shows the combination of individuals, data types, and data/object properties that are used for representing the knowledge graph of the fatigue testing process.



**Figure 4.** Part of FTO graph for representing the “equipment” and “measurement procedure” of fatigue testing. The complete graph can be found in Supporting Information and project repository.<sup>[29]</sup>



**Figure 5.** Part of FTO graph for representing the analysis of fatigue testing data. The blue arrays of A-Box show that the equations for determining fatigue parameters were achieved from the test standard. The complete graph can be found in Supporting Information and project repository.<sup>[29]</sup>

codes of the semantically represented subject–predicate–object triples using the Ontopanel converter. For example, the upper A-Box line of Figure 3 converts to the below TTL codes.

```
fto:FatigueTestExperiment a owl:NamedIndividual,
fto:FatigueTest;
mseo:in_accordance_with fto:TestStandard.
fto:TestStandard a owl:NamedIndividual,
mseo:Regulation;
iof:hasSimpleExpressionValue "ISO 12106:2017"^^xsd:string.
```

Here, fto:FatigueTestExperiment and fto:TestStandard are the individuals, mseo:in\_accordance\_with is an object property, and iof:hasSimpleExpressionValue is a data property. fto:FatigueTest is a class that was created as a subclass of mseo:Experiment>fto:MechanicalTesting classes for representing the fatigue testing process. fto:TestStandard is also an instance for the mseo:Regulation class.

Indeed, the mentioned TTL codes are the OWL representation for the second row of the LCF primary dataset (Figure 1a), which states that the LCF test was performed by the ISO 12106:2017 standard. The other contents of such primary test datasets were also modeled in the same way. The upper part of the A-Box in Figure 3 presents the provenance content of the LCF test, such as the testing date, laboratory, and operator name. The data related to these entities are later (Section 4), provided by linking them with their corresponding values from the primary test dataset (like linking “Versuchs-Datum” with “fto:TestDateIdentifier”).

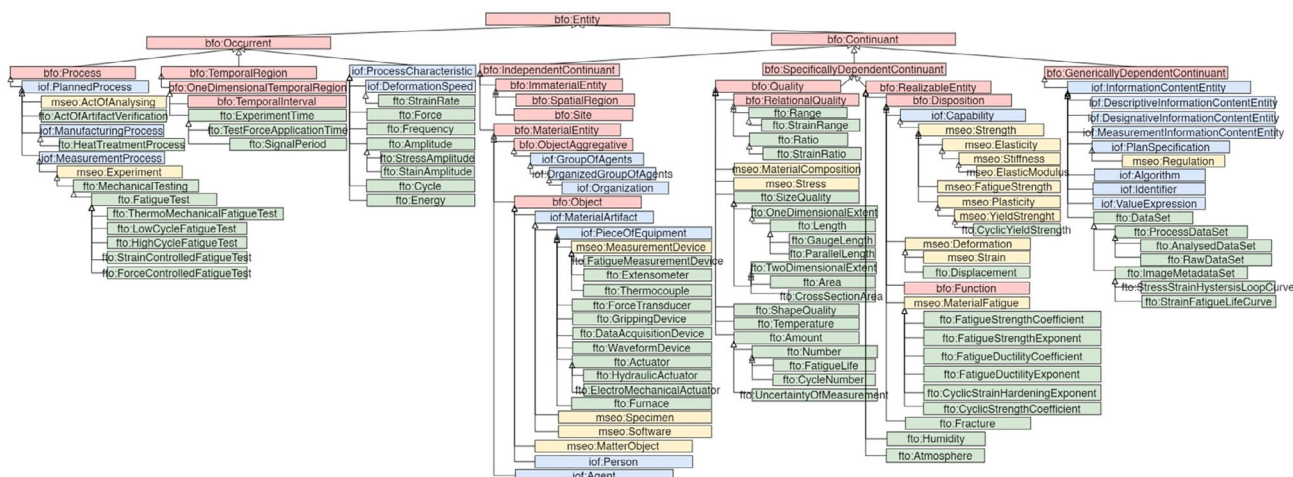
The characteristics of the test pieces as the input of the LCF test are also modeled in the A-Box of Figure 3. These include the test piece identifier, chemical composition, provider, heat treatment, preparation method, cross-section shape, and dimensions. Both the fto:FatigueTestPiece and fto:FracturedTestPiece are instances for the mseo:Specimen class. In the same way, all the modeled features of the test piece as individuals of the knowledge graph were connected to their respective classes in the T-Box of Figure 3.

The specifications of the fatigue test equipment and the detailed LCF measurement process were modeled in the knowledge graph of Figure 4 (the complete graph presented in Figure S1, Supporting Information). The relevant class for expressing the fatigue testing machine has been created as a subclass of mseo:MeasurementDevice. In addition, the components of the fatigue testing equipment (such as force transducer, data acquisition device, waveform device, and actuator) were introduced by the subclasses of iof:PieceOfEquipment, while the device software was represented by the mseo:Software class. It should be mentioned that while many components of the fatigue equipment may be represented in the knowledge graph, FTO only included the amount of information necessary to meet the test standard requirements. Regarding the testing procedure, the measurement steps were also modeled based on the content of the test standard. Here, fto:FatigueTestMeasurement and its steps were introduced as instances of the iof:MeasurementProcess class. Each of these measurement steps had also some qualities (such as test temperature, strain rate, or test time) and their corresponding values should be provided by the content of testing datasets through the data mapping stage (Section 4).

Figure 5 shows part of the fatigue testing knowledge graph which represents the analysis of testing data. The complete version of this graph can also be seen in Figure S2, Supporting

Information. In this case, fto:FatigueTestMeasurementAnalysis and its subsequent steps were represented as instances of the mseo:ActOfAnalysing class. The analysis of fatigue test data is performed in various processes such as calculating stress, strain, fatigue life, cyclic yield strength, fatigue strength coefficient, or cyclic strain-hardening exponent. These analysis procedures contain a few inputs and outputs in addition to certain semantic relationships with the other knowledge graph nodes. For example, in the case of fatigue life determination, the cycle and stress data from the tabular primary dataset are used as inputs, whereas fatigue life (e.g., cycle to fracture) is the process output. Such output data are subsequently linked with the individual values of each quantity. Moreover, the functions and equations that are used for calculating the fatigue test quantities are expressed as the instances of the mseo:Regulation class. As displayed in Figure 5, all these equations are driven from the test standard content.

The knowledge graphs presented in Figure 3–5 were developed in such a generic manner that can model all the entities of the test standard and digitize every content of fatigue test datasets. Collecting all the classes from the T-Boxes of Figure 3–5, the hierarchy of FTO classes was established by adopting the architecture of the BFO, IOF, and MSEO upper-level ontologies (Figure 6). At the top level of this ontology, there are bfo:Entity subclasses (red-colored classes) of bfo:Occurrent and bfo:Continuant. A bfo:Continuant is an entity that survives, endures, or continues to exist throughout time while keeping its identity and has three subclasses: bfo:SpecificallyDependentContinuant, bfo:GenerallyDependentContinuant, and bfo:IndependentContinuant. A bfo:Occurrent is also an entity that unfolds itself in time, and its subclasses of bfo:Processes and bfo:TemporalRegion are highly reused for FTO development. As a midlevel ontology, IOF core provided the entities (e.g., blue-colored classes like iof:MeasurementProcess, iof:ProcessCharacteristic, iof:PieceOfEquipment, and iof:InformationContentEntity) which are needed to model logical and industrial-related aspects of the fatigue ontology. Especially, iof:InformationContentEntity is one of the important FTO classes that is used for mapping the content of test datasets into the knowledge graph. As a domain-level ontology, MSEO also provided various entities (yellow-colored classes) concerning materials and their mechanical testing features, such as the act of analyzing, experiment, measurement device, specimen, stress, strain, strength, material composition, and regulation. Eventually, the green-colored classes are those that are created as the new terms of the FTO. In this case, the FTO classes were developed in four main parts. 1) Classes related to the testing processes that are created on the subclasses of iof:PlannedProcess. Especially the fatigue testing processes were located as the subset of mseo:Experiment>fto:MechanicalTesting. 2) Fatigue equipment classes, such as fto:FatigueMeasurementDevice and fto:DataAcquisitionDevice, are created in the subclasses of iof:PieceOfEquipment. 3) Subclasses of iof:InformationContentEntity (like fto:StrainFatigueLifeCurve) represent the entities related to fatigue testing data. 4) Classes related to the parameters and quantities of fatigue testing are collected in four parts of FTO concerning their theme: subclasses of bfo:RelationalQuality (like fto:StrainRange and fto:FatigueLife), subclasses of bfo:Disposition (like fto:



**Figure 6.** FTO classes' hierarchy concerning BFO, IOF, and MSEO upper-level ontologies (red-, blue-, and yellow-colored class boxes, respectively). The newly added classes with the FTO namespace are marked in green color.

FatigueStrengthCoefficient), subclasses of iof:ProcessCharacteristic (like fto:StrainRate), and subclasses of bfo:TemporalRegion (like fto:SignalPeriod).

Table S1, Supporting Information, provides the full lists of FTO classes in these four parts. It can also be seen in Figure 6 that almost half of the FTO classes are expressed by the higher-level ontologies (red-, blue-, and yellow-colored class boxes). In addition, such higher-level ontologies completely provide all the object properties and data properties needed to express the semantic relations of the knowledge graphs. These findings indicate that the selection and reuse of BFO, IOF, and MSEO upper-level ontologies was quite appropriate for modeling the complex fatigue testing knowledge graph. Reusing these higher-level ontologies promotes the development of knowledge graphs and offers the utilization of well-defined entities and standardizing ontology layout with a global hierarchy, thus ending in higher interoperability of designed knowledge graphs with other ontologies in the fields of materials science and engineering.<sup>[2]</sup>

As previously stated, the graphs in Figure 3–6 were made using the Ontopanel library in the diagrams.net tool and can be converted into RDF files via the Ontopanel convertor. The final steps for completing FTO are done by editing these RDF files with Protégé ontology editor software.<sup>[28]</sup> These steps include importing the higher-level ontologies (for collecting the reused concepts and assuring their correct naming, URI, definition, relation, and hierarchy), checking the hierarchical order of new entities with regard to those of upper-level ontologies, adding explicit definitions, definition references, and further restrictions to newly created entities, completing the basic ontology annotations (like name, version, and creators), reviewing, and testing the developed ontology. **Figure 7** illustrates some FTO features in different Protégé windows. The visualized schema of this ontology is also provided in Figure S3, Supporting Information. FTO is publicly available in the GitLab repository.<sup>[29]</sup>

It is worth noting that the mentioned repository also contains the other versions of FTO that were created by reusing other kinds of higher-level ontologies. FTO V2.x.x for instance, has been developed by reusing PROV ontology (PROVO)<sup>[30]</sup> and

Platform MaterialDigital core ontology (PMDco).<sup>[31]</sup> Figure S4 and S5, Supporting Information, show the graphical views from the T- and A-boxes of FTO V2.x.x, respectively. In this case, the developed ontology modeled all the entities of the fatigue test standard in four parts; subclasses of pmd:Process (to express the test-related processes like fto:FatigueTest), subclasses of pmd:ProcessingNode (about fatigue testing equipment such as fto:FatigueTestingMachine), subclasses of pmd:ValueObject (for modeling the fatigue testing parameters such as fto:FatigueLife), and subclasses of pmd:DigitalEntity like fto:StressStrainHysteresisLoopCurve.

#### 4. ckan.kupferdigital Data Management System for Automatic Data Mapping, RDF Conversion, and Digital Data Storage

This section introduces the ckan.kupferdigital<sup>[32]</sup> data management system which has been designed to streamline and automate the process of data storage, mapping, and conversion. As the leading open-source data portal platform, CKAN<sup>[33]</sup> is an all-inclusive software solution that provides tools for publishing, sharing, finding, and using data (including data storage and robust data application programming interfaces (APIs)) to make data accessible and useful.

The ckan.kupferdigital data management system basically provides three extensions for the CKAN: ckanext-csvwmapandtransform,<sup>[34]</sup> ckanext-csvtocsvw,<sup>[35]</sup> and ckanext-fuseki.<sup>[36]</sup> As part of the ckanext-csvwmapandtransform, the users can prepare the protocols for mapping the datasets into their related knowledge graphs. For this purpose, the user just needs to import a typical test dataset and process knowledge graph into the MapToMethod tool (**Figure 8a**). For the current use case example, we imported the primary dataset for LCF testing of sample G\_21 and FTO.ttl files. In this state, starting the mapping process will be accompanied by the appearance of another window (**Figure 8b**) that allows the user to map all the entities of the test dataset file into the iof:InformationContentEntity entities of the FTO.

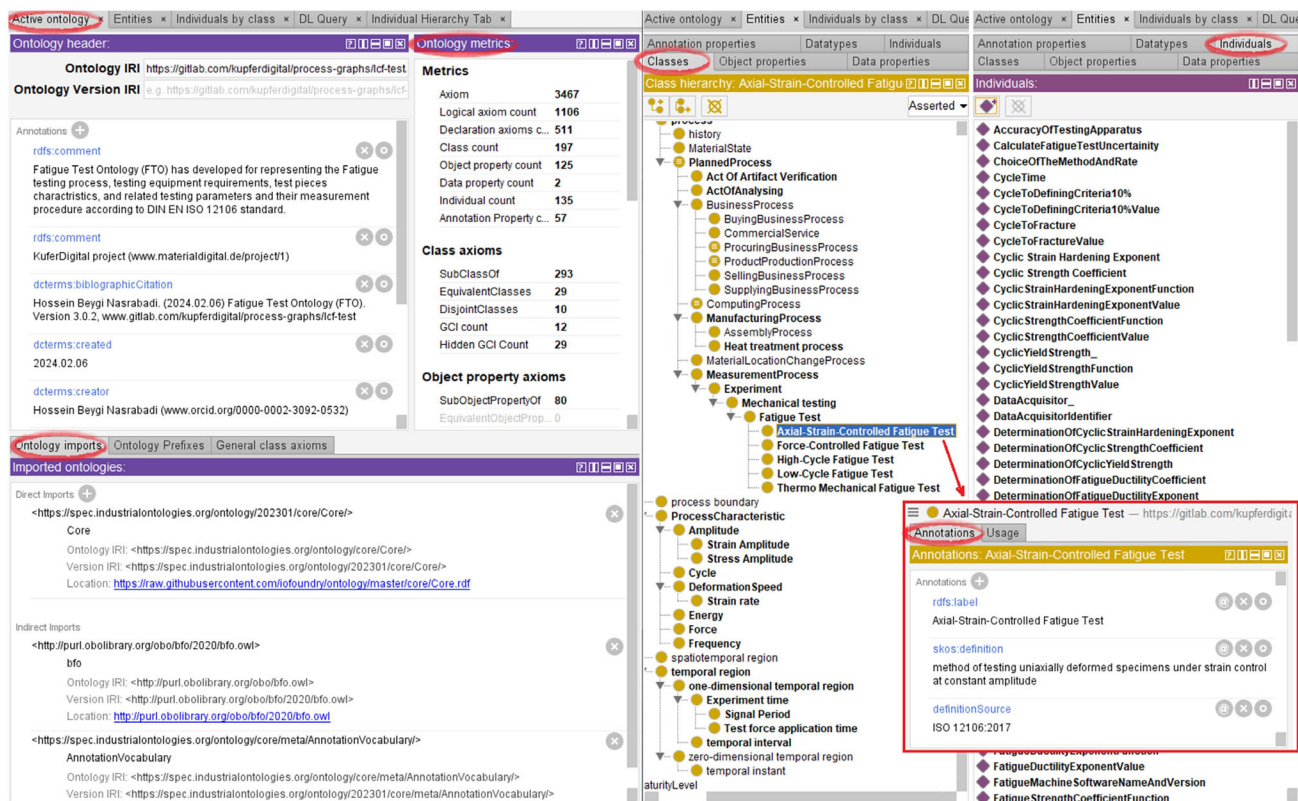


Figure 7. Images of FTO in Protégé showing ontology headers, ontology metrics, imported ontologies, classes hierarchy, annotations, and individuals.

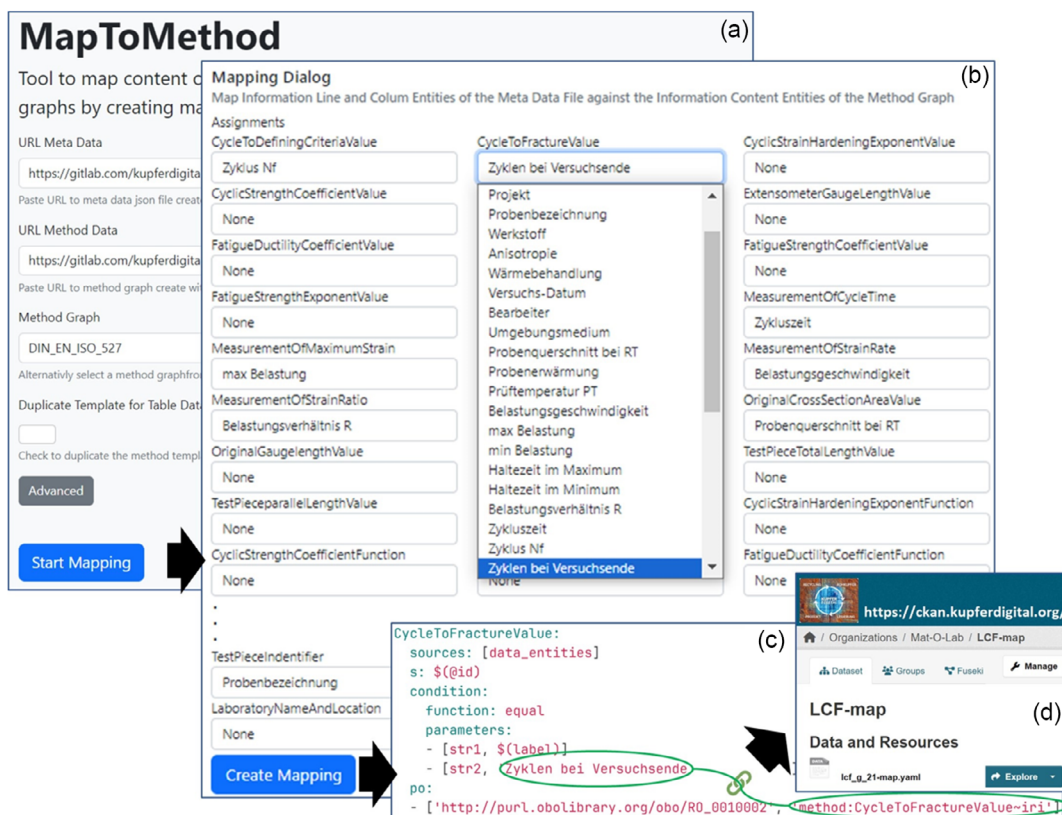
For example, “Zyklen bei Versuchsende” from the test dataset file can be linked with the “CycleToFractureValue” entity of FTO. Similarly, all the other dataset entities are mapped to their associated items of the knowledge graph. Creating the mapping process will then result in a YARRRML file, part of its codes for mapping the CycleToFractureValue data is displayed in Figure 8c. This YARRRML file is eventually stored in the “Process Graph Mappings” group repository of ckan.kupferdigital (Figure 8d).

Based on the provided mapping protocols, ckanext-csvwmapandtransform is capable of automatically converting all the primary dataset files into machine-readable RDF format without the need for any code. For the LCF primary data file of sample G\_21, the automatic conversion of primary testing data in ckan.kupferdigital is illustrated in Figure 9. Here the user just needs to add the new dataset (Figure 9a). A CKAN Dataset is a collection of data resources (such as files), together with a description and other information, at a fixed URL. Datasets are what users see when searching for data. By adding any new primary dataset, ckan.kupferdigital automatically creates JSON and TTL conversions of such data (Figure 9b). In this case, the machine checks the entities of the primary datasets and selects the most matching mapping protocol from the “Process Graph Mappings” group repository. For instance, the machine recognized the “fatigue test mapping” protocol in this use case. However, the users can also follow the “Map&Transform” icon if they would like to use another mapping protocol or even create another mapping protocol.

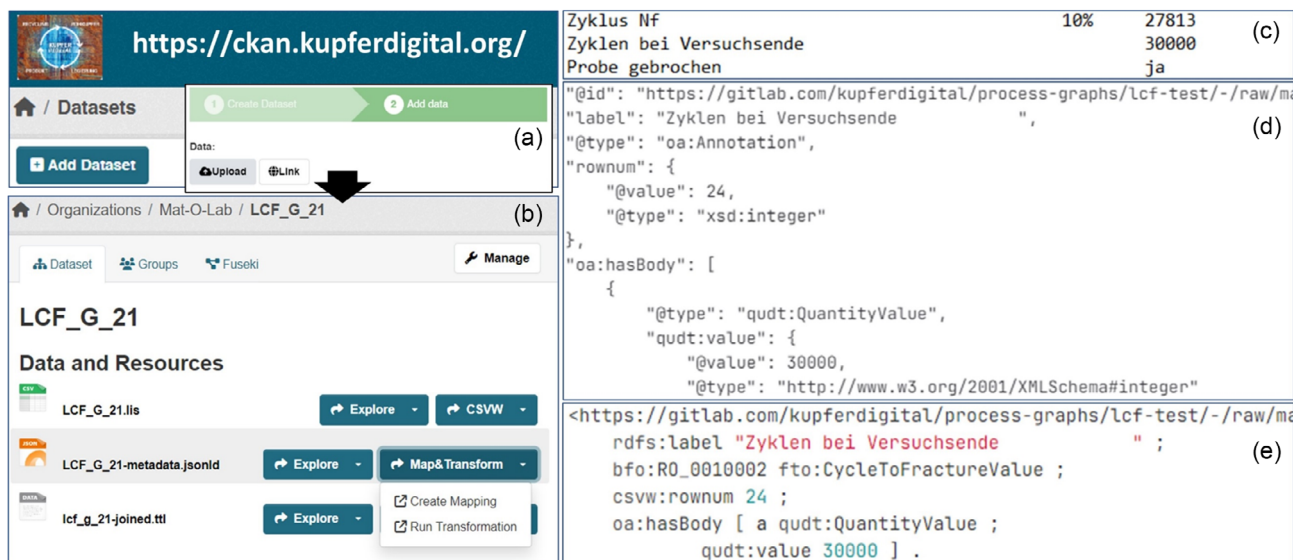
Navigating the “Explore” item, the users can preview and download the files and even plot the figures (like stress–cycle curves) from the primary datasets. Figure 9c–e provides some schema from part of the primary and converted datasets. The primary data of CSV format (Figure 9c) contains various types of metadata which were reported in a human-created structure and German language (e.g., Zyklen bei Versuchsende < space/tab separators >30 000). Based on the ckanext-csvtocsvw extension, all such metadata of primary datasets were then converted into the machine-readable codes of the JSON-LD format (Figure 9d), which represent any metadata within a standardized structure and with a uniform language, unit, and numerical system. Eventually, the metadata of JSON-LD files were mapped into their corresponding entities of knowledge graph (according to the fatigue test mapping protocol) and converted to the TTL formats. For example, it is shown in Figure 9e that “CycleToFractureValue” is linked with “Zyklen bei Versuchsende” and its value is 30 000. Visualizing the resulting TTL files in Figure S6, Supporting Information, it can be seen that all the testing metadata are mapped to the fatigue knowledge graph and intricate relationships and semantics were created between various test metadata.

The outcomes of this section show that ckan.kupferdigital data management system is highly capable of simply and automatically converting the primary test data of heterogeneous structures, languages, and formats into the TTL datasets of standardized content and structure. The RDF-converted fatigue datasets of ckan.kupferdigital have high FAIR scores and can be





**Figure 8.** Data mapping via MapToMethod tool: a) importing data files and knowledge graphs, b) defining the mapping rules, c) creating the YARRRML mapping file, d) storage of mapping protocol in ckan.kupferdigital.



**Figure 9.** A ckan.kupferdigital demonstrator: a) adding new datasets, b) automatic data mapping and RDF conversion, c) preview of typical CSV, d) JSON-LD, and e) TTL datasets.

dynamically queried by complex semantic searches. The third CKAN extension of ckanext-fuseki can be used to seamlessly integrate resources in a CKAN dataset into a dataset in Apache Jena

Fuseki<sup>[37]</sup> triple store. As shown in Figure S7, Supporting Information, by selecting one or more files with compatible triple definitions, a dataset in the triple store named after the CKAN

dataset will be created and the selected resource files uploaded. In this regard, SPARQL (a semantic query language, SQL) queries were utilized to evaluate how the introduced ontology and dataset conversion approaches can result in data retrieval and response to the competency questions. A SPARKLIS<sup>[38]</sup> was integrated to support a human-readable query interface. The individual SPARQL endpoint of the dataset can thus be created by clicking the Query Button in the CKAN Interface (See Figure S7, Supporting Information). Figure S8, Supporting Information, exhibits the results of SPARQL querying the example competency questions from the use case TTL file. In this case, three competency questions were queried: What is the material composition of sample G\_21? Which maximum strain and strain ratio was utilized for the LCF testing of sample G\_21? How much was the number of cycles to failure for sample G\_21? The responses to these queries (CuNi6Sn4, 0.3&-1, and 30 000, respectively) were successfully retrieved by the SPARQL queries, confirming that both the FTO development and ckan.kupferdigital data management system were highly efficient for the data transformation, mapping, storage, and hence digitalization of primary testing datasets.

## 5. Conclusion

Converting the primary material testing data (use case: LCF dataset) to the RDF files and their storage in the open digital repositories was introduced by developing the following two items. 1) FTO: the fatigue testing ontology was designed according to the ISO 12106:2017 standard and modeled the detailed features of fatigue testing specimens and equipment, as well as its testing and data analysis procedure. Designing the knowledge graph based on the testing standards not only provided well-defined domain entities and structured knowledge categorization but also ensured ontology acceptance and applicability in the industry. Utilizing BFO, IOF, and MSEO upper-level ontologies, FTO has been developed in four main parts of the process, equipment, quantities, and testing data. These higher-level ontologies provided most of the FTO modeling entities within a global hierarchy which facilitated the ontology development and its interoperability with other ontologies in the materials science domain. 2) By ckan.kupferdigital: within this data management system and using its MapToMethod tool, the users can prepare the protocols for mapping the datasets into their related knowledge graphs. Based on the provided mappings protocols, the ckan.kupferdigital is capable of automatically converting all the primary datasets into machine-readable RDF files. The data management system enables converting any primary dataset of heterogeneous structures, languages, and formats to OWL-represented and machine-readable data of uniform and standardized structures. Furthermore, ckan.kupferdigital is the open repository for all primary and processed data, so the users can preview and download the files and even plot the figures from the primary datasets.

The use case LCF datasets of several copper alloys were automatically converted to JSON-LD and TTL files within the ckan.kupferdigital. While JSON-LD files are the machine-readable representation of test datasets by the standardized structure, the TTL files are the layout of dataset entities that are mapped into their corresponding entities of knowledge

graph, according to the fatigue test mapping protocol. Such converted and digital datasets have high FAIR scores and are tested by querying the example competency questions. The responses to these queries were successfully retrieved by the SPARQL queries, confirming that both the FTO and ckan.kupferdigital are highly efficient for data transformation, mapping, and RDF conversion.

## 6. Experimental Section

Three grades of industrial copper alloys with compositions of CuNi6Sn4 (alloy G), CuSn8Ni2 (alloy H), and CuZn23Si2.5 (alloy I) were supplied from the fem research institute (fem, Germany). All the samples were produced via the continuous casting method. The CuNi6Sn4 alloy was subjected to a heat treatment process consisting of a heating step at 800 °C for 40 min, followed by water quenching, and another heating step at 410 °C for 3 h. 5–7 fatigue test pieces were prepared from each alloy. The fatigue test pieces had a cylindrical cross section and a total length of 95 mm (parallel length of 17 mm). The dimensions of such test specimens are shown in Figure S9, Supporting Information. The fatigue testing procedures were performed according to the ISO 12106:2017 standard<sup>[5b]</sup> and utilizing a 100 kN electromechanical testing machine (Instron; type 8561; class 1 calibration). An axial extensometer (MTS Systems; type 632.51C-04; 12 mm nominal gauge length; class 1 calibration) was used to measure the strain,  $\epsilon$ . The tests were performed at different strain ranges ( $\epsilon_{\max} - \epsilon_{\min}$ ) while their strain ratio ( $R_{\epsilon} = \epsilon_{\min}/\epsilon_{\max}$ ) and strain rate were  $-1$  and  $10^{-3} \text{ s}^{-1}$ , respectively. Table 1 lists the sample IDs and chosen strain ranges for each fatigue test. All the fatigue tests were carried out in laboratory air and room temperature in the accredited materials testing laboratory of the “Bundesanstalt für Materialforschung und -prüfung.”

## Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

## Acknowledgements

This work was funded by the German Federal Ministry of Education and Research (BMBF) through the project “KupferDigital- Datenökosystem für die digitale Materialentwicklung auf Basis Ontologie-basierter digitaler Repräsentationen von Kupfer und Kupferlegierungen” (grant numbers 13XP5119C and 13XP5119G). The authors would also like to acknowledge Mr. Felix Bauer (fem Research Institute) for providing the materials, Mr. Patrick Uhlemann (BAM) for performing the fatigue testing, and Mr. Gordian Dziwis (Institut für Angewandte Informatik) for assisting with CKAN development.

Open Access funding enabled and organized by Projekt DEAL.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.10820438>, reference number [21].

## Keywords

accessible, CKAN, digitalizations, fatigue testing, findable, interoperable, ontologies, reusable data

Received: March 15, 2024

Revised: June 6, 2024

Published online:

- [1] a) M. Scheffler, M. Aeschlimann, M. Albrecht, T. Bereau, H.-J. Bungartz, C. Felser, M. Greiner, A. Groß, C. T. Koch, K. Kremer, *Nature* **2022**, *604*, 635; b) J. Kimmig, S. Zechel, U. S. Schubert, *Adv. Mater.* **2021**, *33*, 2004940.
- [2] H. Beygi Nasrabadi, T. Hanke, M. Weber, M. Eisenbart, F. Bauer, R. Meissner, G. Dziwis, L. Tikana, Y. Chen, B. Skrotzki, *Comput. Ind.* **2023**, *153*, 104016.
- [3] L. Himanen, A. Geurts, A. S. Foster, P. Rinke, *Adv. Mater.* **2019**, *6*, 1900808.
- [4] A. Jamwal, R. Agrawal, M. Sharma, A. Giallanza, *Appl. Sci.* **2021**, *11*, 5725.
- [5] a) J. Wang, J. Chen, C. Guo, X. Xiao, H. Wang, B. Yang, *Int. J. Fatigue* **2020**, *137*, 105642; b) ISO, ISO 12106:2017 Metallic materials — Fatigue testing — Axial-strain-controlled method **2017**.
- [6] J. Thomas, A. Davis, M. P. Samuel, presented at *Fatigue, Durability, and Fracture Mechanics: Proc. of Fatigue Durability India 2019*, Springer, Bangalore, India **2021**, pp. 267–304.
- [7] a) D. Fensel, in *Ontologies*, Springer, Berlin **2001**, pp. 11–18; b) F. Ekaputra, M. Sabou, E. Serral Asensio, E. Kiesling, S. Biffi, *Open J. Inf. Syst.* **2017**, *4*, 1.
- [8] B. Bayerlein, T. Hanke, T. Muth, J. Riedel, M. Schilling, C. Schweizer, B. Skrotzki, A. Todor, B. Moreno Torres, J. F. Unger, *Adv. Eng. Mater.* **2022**, *24*, 2101176.
- [9] J. Hendler, F. Gandon, D. Allemang, in *Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL*, Morgan & Claypool, San Rafael, CA **2020**.
- [10] T. Ashino, M. Fujita, *Data Sci. J.* **2006**, *5*, 52.
- [11] T. Austin, D. Bullough, D. Gagliardi, D. Leal, M. Loveday, A Guide to the Development and Use of Standards Compliant Data Formats for Engineering Materials Test Data, DE NORMALISATION, COMITÉ EUROPÉEN, and EUROPÄISCHES KOMITEE FÜR NORMUNG **2010**.
- [12] a) N. Hirsch, H. Birkhofer, V. Landersheim, H. Hanselka, U. Gunther, A. Martin, presented at *Int. Design Engineering Technical Conf. and Computers and Information in Engineering Conf.*, Vol. 43277, Brooklyn, USA **2008**, pp. 291–300; b) C. F. Guo, Z. Dong, H. X. Ning, C. D. Wan, *Adv. Mater. Res.* **2014**, *936*, 1894–1899.
- [13] Y. Chen, M. Schilling, P. von Hartrott, H. Beygi Nasrabadi, B. Skrotzki, J. Olbricht, *Integr. Mater. Manuf. Innovation* **2022**, *11*, 545.
- [14] H. Beygi Nasrabadi, B. Skrotzki, *Adv. Sci. Technol.* **2024**, *149*, 33.
- [15] a) J. F. Morgado, E. Ghedini, G. Goldbeck, A. Hashibon, G. J. Schmitz, J. Friis, F. Anne, presented at CEUR Workshop Proc., Heraklion, Greece **2020**; b) M. Schilling, B. Bayerlein, P. von Hartrott, J. Waitelonis, H. Birkholz, P. Dolabella Portella, B. Skrotzki, *Adv. Eng. Mater.* **2024**, 2400138.
- [16] M. A. Bamboat, A. H. Khan, A. Wagan, *Int. J. Emerging Technol.* **2021**, *12*, 25.
- [17] E. Alatrish, *J. Manage. Inf. Syst.* **2013**, *8*, 018.
- [18] S. Mandal, *Library Philos. Pract.* **2022**, *1*.
- [19] N. Haider, F. Hossain, *J. Theor. Appl. Inf. Technol.* **2018**, *96*, 6889.
- [20] D. Calvanese, D. Lanti, T. M. De Farias, A. Mosca, G. Xiao, *Patterns* **2021**, *2*, 100346.
- [21] H. Beygi Nasrabadi, F. Bauer, P. Uhlemann, B. Rehmer, B. Skrotzki, *Zenodo* **2024**, <https://doi.org/10.5281/zenodo.10820438>.
- [22] <https://app.diagrams.net> (accessed: June 2024).
- [23] <http://purl.obolibrary.org/obo/> (accessed: June 2024).
- [24] <https://spec.industrialontologies.org/ontology/core/Core/> (accessed: June 2024).
- [25] [https://github.com/Mat-O-Lab/MSEO/raw/main/domain/util/readable\\_bfo\\_iris.ttl/](https://github.com/Mat-O-Lab/MSEO/raw/main/domain/util/readable_bfo_iris.ttl/) (accessed: June 2024).
- [26] ISO, ISO/IEC 21838-2:2021 Information technology — Top-level ontologies (TLO) — Part 2: Basic Formal Ontology (BFO) **2022**.
- [27] <https://gitlab.com/kupferdigital/process-graphs/lcf-test/-/raw/main/FTO3.0.2.ttl/> (accessed: June 2024).
- [28] <https://protege.stanford.edu/> (accessed: June 2024).
- [29] <https://gitlab.com/kupferdigital/process-graphs/lcf-test/> (accessed: June 2024).
- [30] <https://www.w3.org/TR/prov-o/> (accessed: June 2024).
- [31] <https://github.com/materialdigital/core-ontology/> (accessed: June 2024).
- [32] <https://ckan.kupferdigital.org/> (accessed: June 2024).
- [33] <https://ckan.org/> (accessed: June 2024).
- [34] <https://github.com/Mat-O-Lab/ckanext-csvwmapandtransform> (accessed: June 2024).
- [35] <https://github.com/Mat-O-Lab/ckanext-csvtocsvw> (accessed: June 2024).
- [36] <https://github.com/Mat-O-Lab/ckanext-fuseki> (accessed: June 2024).
- [37] <https://fuseki.kupferdigital.org/> (accessed: June 2024).
- [38] <https://sparklis.kupferdigital.org/> (accessed: June 2024).