Science and Technology of Advanced Materials: Methods

ISSN: 2766-0400 (Online) Journal homepage: www.tandfonline.com/journals/tstm20

A unified model for metallic materials reliability data and its application in NIMS metallic materials database (Kinzoku)

Asahiko Matsuda, Masahiko Demura, Masayoshi Yamazaki, Takuya Kadohira, Toshihiro Ashino, Yoshiyuki Furuya, Kota Sawada & Nobutaka Nishikawa

To cite this article: Asahiko Matsuda, Masahiko Demura, Masayoshi Yamazaki, Takuya Kadohira, Toshihiro Ashino, Yoshiyuki Furuya, Kota Sawada & Nobutaka Nishikawa (2025) A unified model for metallic materials reliability data and its application in NIMS metallic materials database (Kinzoku), Science and Technology of Advanced Materials: Methods, 5:1, 2518745, DOI: 10.1080/27660400.2025.2518745

To link to this article: https://doi.org/10.1080/27660400.2025.2518745

Published online: 14 Jul 2025.

SCIENCE AND TECHNOLOGY OF ADVANCED MATERIALS: METHODS
2025, VOL. 5, NO. 1, 2518745
https://doi.org/10.1080/27660400.2025.2518745

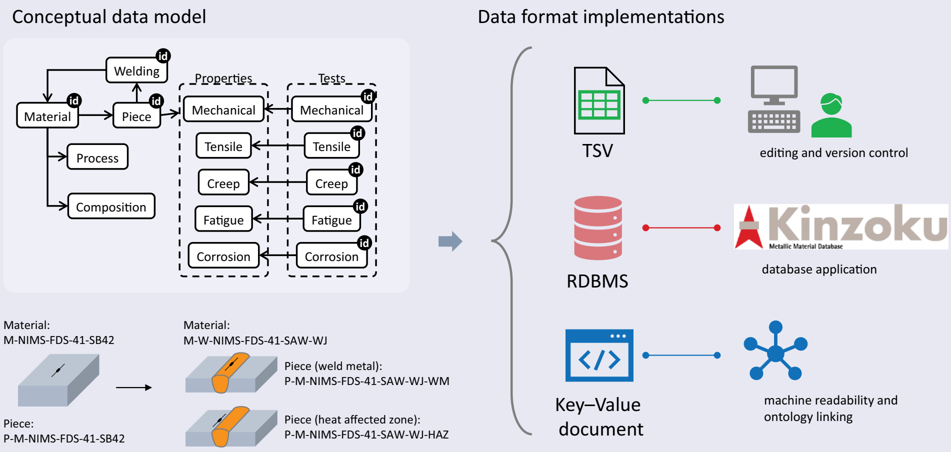
A unified model for metallic materials reliability data and its application in NIMS metallic materials database (Kinzoku)

Asahiko Matsuda [a], Masahiko Demura [a], Masayoshi Yamazaki [a*], Takuya Kadohira [a], Toshihiro Ashino [b], Yoshiyuki Furuya [a,c], Kota Sawada [a,c] and Nobutaka Nishikawa [d]

[a]Research Network and Facility Services Division, National Institute for Materials Science, Tsukuba, Ibaraki, Japan; [b]Faculty of Regional Development Studies, Toyo University, Bunkyo-ku, Tokyo, Japan; [c]Research Center for Structural Materials, National Institute for Materials Science, Tsukuba, Ibaraki, Japan; [d]Mizuho Research & Technologies, Ltd., Chiyoda-ku, Tokyo, Japan

ABSTRACT

A data model to organize metallic materials' reliability properties are presented, along with data format variations to implement the conceptual model. By analysing the structures behind the *Creep Data Sheet* and *Fatigue Data Sheet* series from the National Institute for Materials Science (NIMS), commonalities were identified to establish a standardized list of entities and their relationships. The data model incorporates a multilayered approach that acknowledges both uniform materials and non-uniform materials such as welded joints. Standardized identifier system supports interlinking between the different entities. The model was shown to be applicable to data from various independent initiatives and is effective in supporting datasets for machine learning. The model was implemented in three different formats: spreadsheets, relational database, and key-value document store. Their characteristics are comparatively discussed. Plain-text spreadsheets provide low-barrier editing and are version controllable. Relational database management systems offer data integrity and fast querying, suitable as a backend to a web application; NIMS's Metallic Materials Database (Kinzoku) system was renewed by taking advantage of these characteristics. Key-value document stores can be flexible, highly machine-readable, and self-describing. They also allow for linking with external ontology for heterogeneous data integration. Digitized management of materials reliability data using this model supports storage, federation, and accelerated utilization in data-driven methodologies.



IMPACT STATEMENT

The unified data model and its various formats allow description of complex specimens, provides interoperability across different data sources, streamlines data integration for databases, and facilitates machine learning applications.

ARTICLE HISTORY
Received 23 April 2025
Revised 19 May 2025
Accepted 5 June 2025

KEYWORDS
Metallic materials; structural materials; materials database; data integration; data format; data model; creep; fatigue; materials reliability

1. Introduction

Reliability of metallic materials is a vital issue in our society, as these materials must carry crucial load under elevated temperatures and/or large pressure differences in power plants, boilers, turbines, etc. To assist research and development of advanced

CONTACT Masahiko Demura DEMURA.Masahiko@nims.go.jp Research Network and Facility Services Division, National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan

*Retired.

Supplemental data for this article is available on NIMS Materials Data Repository at https://doi.org/10.48505/nims.5449.

materials, data-driven approaches to extract knowledge from large datasets using data analysis methodologies have gained traction [1–4]. Digital platforms focusing on data-driven research are expected to accelerate the integration of machine learning, calculations, and experimental data, and may be referred to as materials discovery platforms or materials acceleration platforms [1,2,5–8]. On the other hand, decades before the rise of these digital platforms, various initiatives have been collecting valuable data on metallic materials reliability. Notable publications and initiatives include the *ASTM Data Series* [9–11], Creep Committee of the Iron and Steel Institute of Japan (ISIJ) [12–16], the European Creep Collaborative Committee [17], and the structural materials data sheets by National Institute for Materials Science (NIMS) [18–23]. These data were initially compiled as printed publications, but digitization and integration of data from these sources provide great benefits, as the effectiveness of data-driven approaches depends on the amount and diversity of data available for machine-learning model training and analysis [1,3,5]. The data sources were designed and collected independently from one initiative to another and hence their data are reported in various layouts and formats. However, these data sheets inherently have a well-organized conceptual data model. It is helpful to distinguish between designing a conceptual data model and more concrete data format implementations [24–26]. Data modelling in this context includes identifying what kinds of entity classes exist for the subject matter, designing the relationships between the entities and their properties, and establishing a system for identifiers to link the entities. The data sheet publications may not specify the data model and schema explicitly enough for digital databases as to be programmatically unambiguous, but they exist implicitly and are reflected in the how the data sheets are organized. Therefore, in this paper, the structures of *NIMS Creep Data Sheet* (CDS) and *NIMS Fatigue Data Sheet* (FDS) [18–21] are analysed to extract their commonalities and a data model for digital representation is defined. Compatibility of the model with data from other initiatives and applicability in machine learning are discussed. Finally, we will describe several different implementation formats for this data model: spreadsheets, relational database, and key-value document store. The advantages and use case scenarios for each format are comparatively discussed.

2. Data Sheet structure analysis

2.1. NIMS Creep Data Sheet

The Creep Data Sheet Project was established in 1966 by the National Research Institute for Metals (NRIM – now NIMS through merger), initially advised by the ISIJ Creep Committee [18,27]. The data sheets were originally published in paper form, and are now published online on NIMS's MatNavi database service [20,28–30] (registration required). 60 data sheets have been published as of April 2025. The creep tests run continuously for years, so the published data sheets receive updates, indicated by revision letters after the data sheet number: A for the first revision and B for the second revision.

A typical CDS structure is summarized in Figure 1. The creep tests are performed as specified in JIS Z 2271. Each material is assigned a three-letter NRIM/NIMS reference code, which is assigned to be uniquely identifiable across all CDS. The details of the subject materials (reference code, type of melting, ingot size, etc.) are given first, followed by the table of chemical composition for each reference code, short-time tensile properties for various test temperatures, and creep rupture data for various test

Common among typical Creep Data Sheets:

1. Details of materials

Columns	***Example entry***
Reference code	*MBB*
Type of melting	*BEA (basic electric arc)*
Size of ingot (kg or t)	*4.85*
Deoxidation process	*Si-killed*
Product form	*Tube*
Dimension (mm)	*50.8 OD 8 WT 5000 L*
Processing/thermal history	*(omitted)*
Grain size	*7*
Hardness	*83*
Non-metallic inclusion (%)	

2. Table of chemical composition

Requirements for each element *(a column for each element)*
a row for each reference code:
Chemical composition (mass %) *(a column for each element)*

3. Short-time tensile properties

a row for each reference code:
a column for each test temperature:
0.2% proof stress
Tensile strength
Elongation (%)
Reduction of area (%)

Data plots and regression lines

4. Creep rupture data

a table for each test temperature:

a row for each reference code:
a column for each stress:
Time to rupture (hours)
Elongation (%)
Reduction of area (%)

Data plots and regression lines

In some Creep Data Sheets:

Summary of rupture strength
- Rupture curve predicted from data points
- Temperature dependence of 0.2% proof stress, tensile strength, and creep rupture strength

Different tabulation of creep data

Data plots
- Stress versus time
- Time to rupture versus minimum creep rate

Microstructure photographs

Profile photographs of specimens

Supplementary data
- Specimen sampling
- Test methods
- Analysis and representation methods

Figure 1. Typical contents of a NIMS Creep Data Sheet (CDS).

temperatures. Many data sheets have additional data such as rupture strengths, data plots, and photographs. The data are tabulated, with short-time tensile properties and creep rupture data plotted as figures in addition to the tables.

A sheet's first table defines the reference codes and describes the basic processing parameters. In addition, some microstructural features and basic material properties, such as grain size and hardness, are included as well. Chemical compositions are also important basic property of the materials, but these are better expressed separately, and thus CDS places them in their separate section. For a given material reference code, the measured properties take on different dimensions. Hardness is measured in room temperature and is always a scalar for a given reference code, but the short-time tensile properties are functions of the test temperature. Thus, while the hardness can be written on the same table with the specimen processing details, the tensile properties must be given in a separate table, and this in fact is how the CDS is structured. Creep rupture data require another dimension, as the time to rupture, elongation, and reduction of area are respectively functions of both test temperature and applied stress; hence, they are represented as a series of multiple tables on the data sheets. In other words, the nested structure of rows, columns, and tables as listed in Figure 1 is indicative of the dimensionality of the properties.

CDS Nos. 32A and 45A focus on welded joints, which welds the base metals in various configurations, and assigns NIMS reference codes to the welded metal and the joints [31,32]. The regions that have melted by welding, the regions that experienced heat, and the regions far away from the joint will exhibit different properties, and thus a data model to express this complexity will require an additional layer between the material and the property. This will be covered in the later section where we discuss the data model.

2.2. NIMS Fatigue Data Sheet

The first NRIM Fatigue Data Sheet was published in 1978 [33], and since then, a total of 136 data sheets have been published as of February 2025. These data sheets are now published online [21,28,29]. They contain fundamental fatigue properties of structural materials at room temperature and at high temperatures. The materials are made in Japan (JIS-grade steels, aluminium alloys, titanium alloys, etc.) and tested in NIMS, which include high-cycle, low-cycle, and gigacycle fatigue tests [19]. The data sheets are typically sectioned as shown in Figure 2, starting from the Material section (processing details, chemical composition and heat treatment conditions) and Welding section (present when applicable), followed by Mechanical Properties and Fatigue Properties sections.

Each heat is identified within the data sheet by a single-letter local identifier, unique within the data sheet. Processing details and the following chemical composition tables are structured similarly to CDS, while FDS frequently distinguishes between product analysis and ladle analysis.

Mechanical properties are given in a separate section from the processing details, with a row for each heat and a column for each property such as the various tensile properties, Charpy impact value, and hardness.

In the Fatigue Properties section, a table for the test conditions are given first, followed by the test data subsection. Test data are organized as a series of tables, with a table for each test condition, a column group for each heat, and a row for each number. Stress or strain conditions and the number of cycles to failure are reported on each row. *S*–*N* diagrams (stress vs. number of cycles to failure) are shown.

3. Data model

3.1. Materials and pieces

A trivial data model for a typical materials database would have a list of properties that references

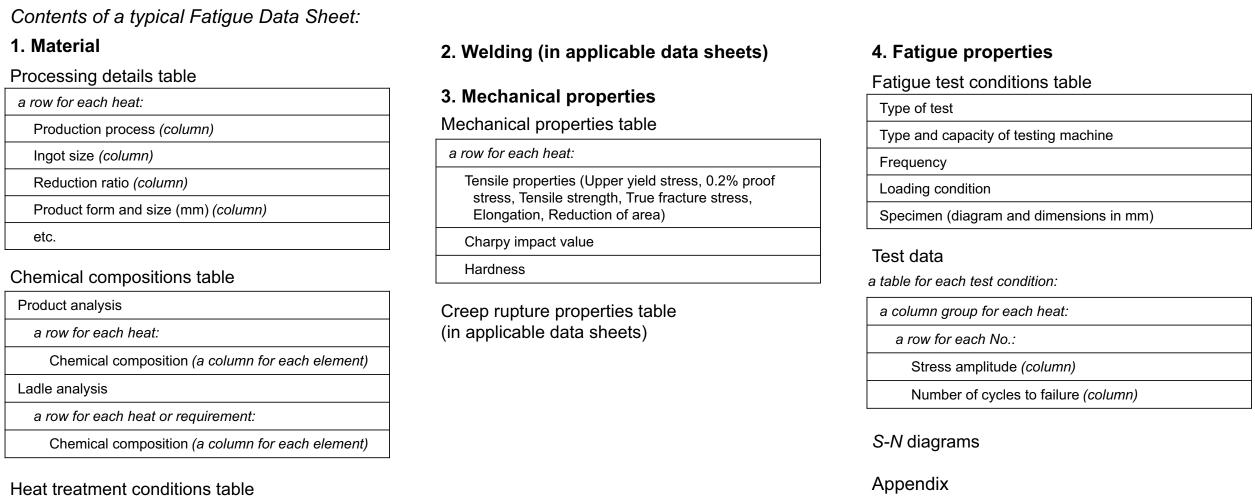


Figure 2. Typical contents of a NIMS Fatigue Data Sheet (FDS).

materials, as illustrated in Figure 3(a). However, this simple model implicitly assumes uniformity throughout the specimen being studied. Some of our data focus on welded joints, where regions with different properties – e.g. base metal, heat-affected zone, and weld metal – coexist on a single welded joint (FDS No. 41 for example [34]). Therefore, it becomes necessary to define multiple 'pieces' from the 'material' in question and differentiate between the two. The concept of 'pieces' is not explicitly spelled out in the original CDS/FDS and is defined in this data model as tangible parts within the material in question. It is useful to introduce this distinction to materials other than welded joints as well, as many data sheets deal with multiple pieces from a material due to reasons such as:

- different heat treatments (FDS 24–26, 29, 30, etc.),
- specimen processing (polishing) (FDS 37),
- solution treatments (FDS 54),
- longitudinal or transverse direction (FDS 124, 126, 127), and
- specimen shape variations (FDS 47, 57, 65, 117).

Thus, we arrive at a model illustrated in Figure 3(b). For materials that are not welded joints, we define the 'material' and take 'pieces' from it. Whereas a 'material' is tied to its compositions, standard, and symbol of grade, a 'piece' is tied to properties such as thickness and heat treatment condition. Also, a definition of a piece references the material from which it was taken from. For welded joints, the base materials are defined, followed by the pieces tested for the material (each piece references a material), then the welded joint (references the pieces and weld metal), which is recorded as another entry as a material (references the welding), from which pieces are taken again (references the welded material), which is finally tied to the specific properties.

3.2. Tests and properties

There are several categories of material properties for a given material and its piece: tensile properties, creep properties, elastic modulus, and fatigue properties, which are measured by their respective tests: tensile tests, creep tests, elastic modulus tests, and fatigue tests. For these, corresponding entities for the respective material properties and their tests are introduced. In database design, it is common to normalize the structure in order to reduce data redundancy and increase data integrity [24]. Experimental conditions that are common throughout each test should be normalized as database properties of their respective tests, such as test standards (e.g. JIS Z 2271 for creep tests, JIS G 0567 for some tensile tests), testing machine, and loading conditions (load waveform, frequency, etc.). On the other hand, some experimental conditions are left unnormalized, a prime example being the test temperature. This quantity serves as an independent variable for the experiment and thus is modelled along with the corresponding material properties. For example, database properties for the entity 'creep property' contain the test temperature, test stress, time to rupture, elongation, and reduction of area. This structure also makes editing the data more straightforward.

3.3. Entities and identification system

Table 1 lists the main entity types in our data model. These data entities are assigned unique identifiers, and each ID type follow a defined format to ensure traceability. For instance, there is a fatigue data sheet on the properties of SB 42 carbon steel, which is numbered FDS No. 41 and hence is identified in our data as 'D-NIMS-FDS-41' [34]. This data sheet reports on data for the uniform base metal as well as butt-welded joint. The uniform SB 42 base metal is given the material identifier 'M-NIMS-FDS-41-SB42'. A piece taken from this base metal is identified as 'P-M-NIMS-FDS-41-SB42', incorporating the material identifier. For the butt-welded joint, the submerged arc welding (SAW) process in this data sheet is identified as 'W-NIMS-FDS-41-SAW'. The welded joint created by this process is given the material identifier 'M-W-NIMS-FDS-41-SAW-WJ' by incorporating the welding ID and adding the material prefix 'M-' and the welded joint suffix '-WJ'. Different pieces are taken from this welded

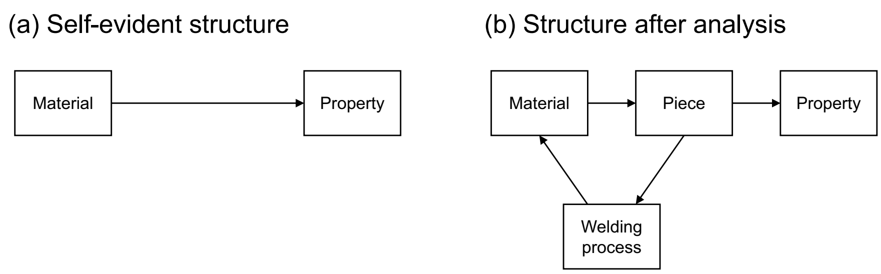


Figure 3. Data model groundwork for describing materials and their properties, with the layer for pieces included to allow for welded joints and several other circumstances.

joint: 'P-M-W-NIMS-FDS-41-SAW-WJ-WM' from the weld metal (melted part) and 'P-M-W-NIMS-FDS-41-SAW-WJ-HAZ' from the heat-affected zone. There is another weld metal piece that has undergone post weld heat treatment (PWHT), identified as 'P-M-W-NIMS-FDS-41-SAW-WJ-WM-HT'. Uniaxial fatigue tests are performed on these pieces, whose test conditions are identified as e.g. 'T-F-NIMS-FDS-41-UNI-1', the letter F indicating fatigue property and UNI indicating uniaxial test type. Tensile tests are also performed, its test condition being identified as 'T-T-NIMS-FDS-41-1' (second T for tensile).

The identifier formats for all the relevant entities are summarized in Table 1. In the database, these IDs can be used for referencing the entities from other types of entities as foreign keys. Process, heat treatment, material composition, and material properties do not need to be referred to from other entities in this model and are not part of the ID system. There are ten types of material properties defined:

- Creep property
- Stress relaxation
- Tensile property
- Mechanical property
- Elastic modulus
- Fatigue property:
  - Load-controlled
  - Fatigue strength
  - Strain-controlled incremental step test
  - Strain-controlled constant amplitude test
  - Fatigue crack propagation

For each of these property types, there are multiple quantities to be recorded. For creep property, these include test temperature, test stress, time to rupture, elongation, and reduction of area. For load-controlled fatigue property, these include stress amplitude and number of cycles to failure. See supplemental data for the other properties and further details.

3.4. Model applicability

Our earlier analysis was based mostly on NIMS CDS and FDS, but we were able to successfully digitize data from the following sources also using our data model.

- Tensile and creep-rupture properties in *ASTM Data Series* [9–11]
- *Report on the Mechanical Properties of Metals at Elevated Temperatures* by ISIJ [12–16]
- Database System for Pressure Vessel Materials by Japan Science and Technology Corporation (JST) and the Materials Division, Japan Pressure Vessel Research Council (JPVRC) [35,36]
- *Corrosion Data Sheet* and *Space Use Materials Strength Data Sheet* by NIMS [22,23]

These sources combined represent a considerable amount of data (Table 2), and we believe that our data model can serve as the common foundation in recording reliability data for metallic materials under wide circumstances. We also find commonalities with the format used in a recent literature-derived database, FatigueData-CMA2022 [37], which features a hierarchical tree branching from the data sources (articles) to multiple data structs of materials, processing, testing, mechanical properties, and fatigue. These can be mapped to corresponding entities in our data model so that a comprehensive dataset can be compiled.

Datasets organized using a consistent data model allows large amount of data to be processed for machine learning. By compiling a dataset from CDS in a data format supported by the data model described here, Sakurai et al. [38,39] constructed prediction models that can predict creep rupture time from chemical composition, test temperature, and stress. In preparing their dataset, they needed to manage data which shared the same chemical composition but were different in terms of creep

Table 1. Entity types with IDs and their formats.

Entity type	ID format	Examples	Note
Data source	D-[org/source]	D-NIMS-CDS-1 D-NIMS-FDS-41	[org/source] = '[organization]-[source]-[sheet no.]'
Material	M-[org/source]-[material code]	M-NIMS-CDS-1-STBA22-MBB M-NIMS-FDS-41-SB42	
Piece (for uniform material)	P-[Material ID]	P-M-NIMS-CDS-1-STBA22-MBB P-M-NIMS-FDS-41-SB42	
Welding	W-[org/source]-[welding process]	W-NIMS-FDS-41-SAW	Welding processes: GMA = gas metal arc welding, MAC = manual arc welding covered electrode, SAW = submerged arc welding.
Material (welded)	M-[Welding ID]-[zone]	M-W-NIMS-FDS-41-SAW-WJ	WJ = welded joint
Piece (for nonuniform material)	P-[Material ID]-[piece code]	P-M-W-NIMS-FDS-41-SAW-WJ-WM P-M-W-NIMS-FDS-41-SAW-WJ-HAZ	Piece codes: HAZ = heat-affected zone, WM = weld metal, WM-HT = weld metal after post weld heat treatment
Test condition	T-[property code]-[org/source]-[test type/#]	T-C-NIMS-CDS-C-1 T-F-NIMS-FDS-41-UNI-1	Property codes: C = creep, *T* = tensile, EM = elastic modulus, F = fatigue. Test types: C = creep, CR = creep rupture, A = axial, RB = rotating bending, TOR = torsion, UNI = uniaxial.

Table 2. Numbers of materials, welding entities, and property data in our spreadsheet representation. Original data sources are cited in text.

Source	Materials	Welding	Tensile	Creep	Fatigue
NIMS CDS	431	22	4136	10853	0
NIMS FDS	632	93	1069	210	39941
ASTM	536	0	1456	1599	0
ISIJ	1109	168	4521	10052	0
JPVRC	192	0	570	961	0
NIMS SDS	60	26	99	36	7917

conditions. In doing so, using this model's piece IDs to extract data was effective in preventing missed or duplicate data. Specifically, they employed stratified sampling using the piece IDs for making the test dataset, which was made possible by this model. Most recently, Sakurai et al. [40] reported on a collaborative prediction model development using federated learning, a confidentiality-preserving machine learning technique that enables prediction model training without direct data sharing between parties. The eight participating organizations were able to build a global prediction model by sharing only limited information such as local model parameters, while keeping their data localized and confidential. Their global model was shown to be more robust, performing better across all datasets than a single institution's local model. Since the participating organizations did not have direct access to each other's data, it was crucial to work on a common data model and format to organize their heterogeneous data. The data format used for this federated learning project was based on the earlier work [38,39] and further developed through coordination between the organizations.

4. Data formats

The data model described in this paper can be implemented in several different formats. We will cover three basic variations in this paper: spreadsheets, relational model, and document-oriented key-value store.

4.1. Spreadsheets

Data managed as simple spreadsheets (Figure 4) benefit from high browsability and ease of editing. This makes it suitable as the format for day-to-day data management. On the other hand, spreadsheets do not have sufficient ability to self-describe the structure of the data and the relations between the constituent tables; an external document is necessary to describe the full picture of the data. This paper and the supplemental data serve as such documents. We will see later that the spreadsheets can serve as the source format to be converted into other formats featuring more comprehensive self-description, data validation, and/or semantic linking.

Each of the entity types we have defined in our previous section (materials, pieces, etc.) has a fixed set of properties to be recorded. Therefore, each type will form a table, wherein instances of the entities are listed as rows and the properties for each instance are organized into columns. All of the information for one instance of the entity should be recorded on the same row, typically starting with the ID of the entity instance as we have defined in Table 1, followed by its name and properties. This can be done for each property as long as the property can be expressed as a single string or scalar value, but whenever the property calls for more dimensions, it must be split into a separate table. Entries in the separate table can reference entries in other tables using the entity IDs (Figure 5). Additionally, we have decided to split some properties into their separate table on a case-by-case basis. For example, atomic composition of a material entity consists of more than forty columns, each column holding a mass percentage of an element. Considering the conceptual grouping of this large column group, these have been split into a separate 'composition' table, which also helps to keep the material table from becoming too wide.

Each table starts with header rows to indicate what the column contains. In our format, we have a row for the column names and a row for the units. While common spreadsheet formats such as Microsoft Excel workbooks can be used, we adopted plain-text tab-separated values (TSV). This is because spreadsheet editors may truncate the trailing zeroes after decimal points in our data if it is loaded as numbers. We prefer to honour the number of significant digits in our data, so editors which treat the data as text (e.g. EmEditor, Visual Studio Code with CSV extensions, or SmoothCSV) are preferrable. Each table is separated into its own TSV file – one table per file. The files' management can be either centralized or distributed, using data comparison utility 'diff' and version control software such as Git, taking advantage of their plain-text nature.

4.2. Relational database

Our data model was also implemented using a relational database management system (RDBMS), which has the benefits of better data integrity and

[Material]

Material ID	Datasource ID	Material	Nominal Composition	Material Standard	Material Category
M-NIMS-CDS-1-STBA22-MBB	D-NIMS-CDS-1	STBA22	1Cr-0.5Mo	JIS G 3462	Low alloy steel
M-NIMS-FDS-41-SB42	D-NIMS-FDS-41	SB42	0.16C-0.80Mn	JIS G 3103	Carbon steel
M-W-NIMS-FDS-41-SAW-WJ	D-NIMS-FDS-41				

[Composition]

Material ID	Analysis	C	Si	Mn	P	S	Ni	Cr	…
		mass%	mass%	mass%	mass%	mass%	mass%	mass%	…
M-NIMS-CDS-1-STBA22-MBB	Product analysis	0.12	0.36	0.52	0.009	0.003	0.089	0.97	…
M-NIMS-FDS-41-SB42	Ladle analysis	0.16	0.19	0.80	0.016	0.007	0.01	0.02	…

[Piece]

Material ID	Piece ID	Welding ID	Zone	Crack Propagation Zone
M-NIMS-CDS-1-STBA22-MBB	**P-M-NIMS-CDS-1-STBA22-MBB**			
M-NIMS-FDS-41-SB42	**P-M-NIMS-FDS-41-SB42**		Base metal	
M-W-NIMS-FDS-41-SAW-WJ	**P-M-NIMS-FDS-41-SAW-WJ-HAZ**	W-NIMS-FDS-41-SAW	Butt welded joints	Heat-affected zone

[Heat Treatment]

Piece ID	Heat Treatment 1 Temperature	HT1 Time1	HT1 Condition	Heat Treatment 2 Temperature	HT2 Time1	HT2 Condition	Heat Treatment 3 Temperature	…
	C	h		C	h		C	…
P-M-NIMS-CDS-1-STBA22-MBB	910C	0.167	Air cooling	670C	1.167	Air cooling		…

[Welding]

Welding ID	BM1 Piece ID	BM2 Piece ID	Welding Method Type	Shape of Welded Joint	Weld Length	Welding Current	Welding Speed	Welding Voltage
						A	cm/min	V
W-NIMS-FDS-41-SAW	P-M-NIMS-FDS-41-SB42		SAW	Butt welded joint	500	670	32	36

Creep

[Creep Test]

Test ID	Test Standard	Test Method
T-C-NIMS-CDS-CR-1	JIS Z 2271	Creep rupture test
T-C-NIMS-CDS-C-1	JIS Z 2271	Creep test

[Creep Property]

Piece ID	Test ID	Test Temperature	Test Stress	Time to Rupture	Test Elapsed Time	Elongation	Reduction of Area
		C	MPa	h	h	%	%
P-M-NIMS-CDS-1-STBA22-MBB	T-C-NIMS-CDS-CR-1	500	373	***		***	***
P-M-NIMS-CDS-1-STBA22-MBB	T-C-NIMS-CDS-CR-1	500	333	***		***	***
…	…	…	…	…	…	…	…
P-M-NIMS-CDS-1-STBA22-MBB	T-C-NIMS-CDS-CR-1	550	265	***		***	***
P-M-NIMS-CDS-1-STBA22-MBB	T-C-NIMS-CDS-CR-1	550	216	***		***	***

Tensile

[Tensile Test]

Test ID	Test Standard
T-T-NIMS-FDS-41-1	JIS Z 2201
T-T-NIMS-FDS-41-2	JIS Z 3111
T-T-NIMS-FDS-41-3	JIS Z 3121

[Tensile Property]

Piece ID	Zone	Test ID	Test Temperature	Upper Yield Stress	0.2% Proof Stress	Tensile Strength	Elonga-tion	Reduction of Area
			C	MPa	MPa	MPa	%	%
P-M-NIMS-CDS-1-STBA22-MBB			RT		***	***	***	***
P-M-NIMS-CDS-1-STBA22-MBB			100		***	***	***	***
P-M-NIMS-CDS-1-STBA22-MBB			200		***	***	***	***
P-M-NIMS-FDS-41-SB42		T-T-NIMS-FDS-41-1		***		***	***	
P-M-W-NIMS-FDS-41-SAW-WJ	welded joints	T-T-NIMS-FDS-41-3				***		

Fatigue

[Fatigue Test]

Test ID	Datasource ID	Type of test	Testing Machine	Loading Condition	Stress Ratio	Waveform	Test Temperature	Frequency
T-F-NIMS-FDS-1-RB-1	D-NIMS-FDS-1	Rotation bending	4-point loading 100 N-m	Load control			RT	50
T-F-NIMS-FDS-1-TOR-1	D-NIMS-FDS-1	Torsion	Rotating eccentric mass 50 N-m	Load control			RT	33
T-F-NIMS-FDS-41-UNI-2	D-NIMS-FDS-41	Uniaxial	Servohydarulic 0.4 MN	Load control	0	Sinusoidal	RT	1-60

[Fatigue Property (Load-Controlled)]

Piece ID	Test ID	Test Temperature	Frequency	Stress Amplitude	Number of Cycles to Failure	Number of Cycles Elapsed
		C	Hz	MPa		
P-M-NIMS-FDS-1-S25C-A	T-F-NIMS-FDS-1-RB-1	RT		290	***	***

Figure 4. Examples of entries from some of the spreadsheet tables. The topmost lines in square brackets are not part of the actual table. Header rows and the primary identifier for the entity type are in bold. Only three property tables are shown here for brevity. See supplemental data for all tables and their columns.

powerful querying capabilities using Structured Query Language (SQL) [24]. However, creating and updating database records using SQL in laboratory settings can be tedious and impractical. Therefore, we developed a batch program which takes the TSV files as the source data and imports them into our PostgreSQL implementation. A part of its entity–relationship diagram is illustrated in Figure 6. Each of the TSV files in our spreadsheet format corresponds to a table in the relational database. The IDs defined in Table 1 can serve as the primary keys (PK) and foreign keys (FK) for those entities.

Columns for the PostgreSQL tables mostly follow those of the corresponding TSV files, and hence loading the database from the source TSV is mostly a straightforward process of simply copying the data

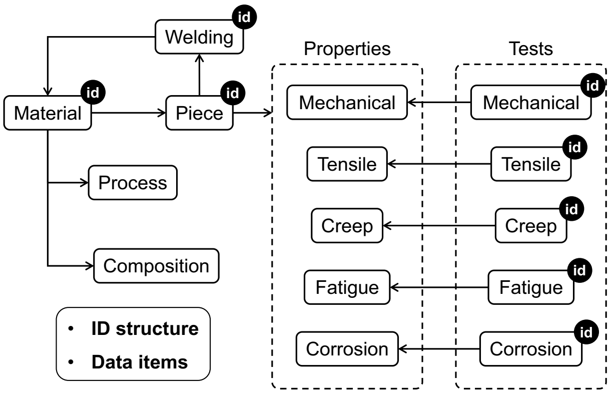


Figure 5. External table referencing through entity IDs.

values. For the columns expecting numerical values such as the composition table and material properties, data are stored in two columns of different data types: one for text and the other for numeric. If the TSV contains a numeric value such as '500', the value is simply stored redundantly in both the text and numeric columns. For values such as '0.80', the numeric column stores the value 0.8, while the text column will retain the trailing zero to preserve the number of significant digits in the data source. Test temperatures are frequently reported as just 'RT' in the data source, standing for room temperature. The text column stores the string 'RT' verbatim, while the numeric column stores 25 (°C). Other rules are listed in Table 3. These rules were incorporated into the aforementioned batch program, which takes the TSV files as the input, applies the rules, checks for ID uniqueness, scans for any other errors against the database schema, and converts the data into a format for PostgreSQL loading.

Using this new database as a backend, we renewed our web-based database application called Metallic Materials Database (Kinzoku) and released a new version in 2023 [41] (registration required). The application provides the graphical user interface for exploring the database, capable of searching, viewing, and dynamically visualizing the data as a chart (Figure 7). This application is the successor to the application previously provided under the same name, which was released in 2010 [42,43]. The older version was also capable of searching and presenting data across the NIMS data sheets, but the underlying database system contained several data models coexisting side-by-side: one for CDS, one for FDS, etc. A key point in this paper's new implementation is that this has been replaced with a single consistent schema unifying all data sources at the data modelling level.

4.3. *Key-value document store*

Key-value databases are a type of non-relational database that store data as a collection of key-value pairs.

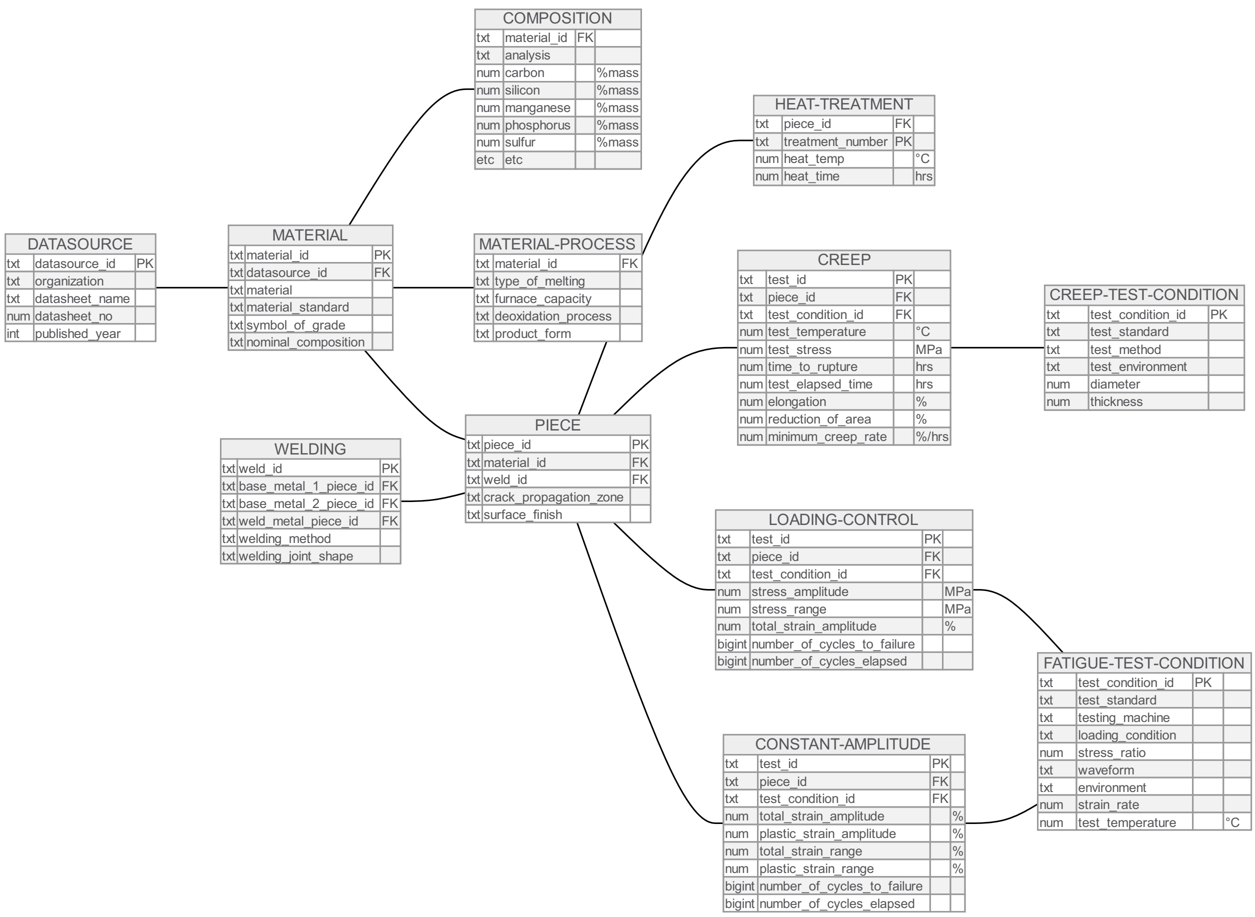


Figure 6. Part of the entity–relationship diagram for the PostgreSQL implementation of our data model. The fourth column in some tables indicate either descriptions of the properties or units of measure.

Table 3. Examples of value conversion rules in the RDBMS data loader batch program.

Type of value	TSV entry example	Numeric conversion
Temperature (°C)	RT	25
Composition (mass%)	Trace	0
Composition (mass%)	Balance	100 – sum of all other elements
Range	12.0–15.0	13.5

Many key-value formats also allow for nested elements by containing such collection as the value or by having hierarchical structure. We implemented our data model using Extensible Markup Language (XML), but it is possible to use other formats (e.g. JSON) or dedicated key-value store solutions. The XML and JSON approaches are known as document-oriented databases, where the information for a given entity is stored in its own document as opposed to being spread across multiple tables like in RDBMS.

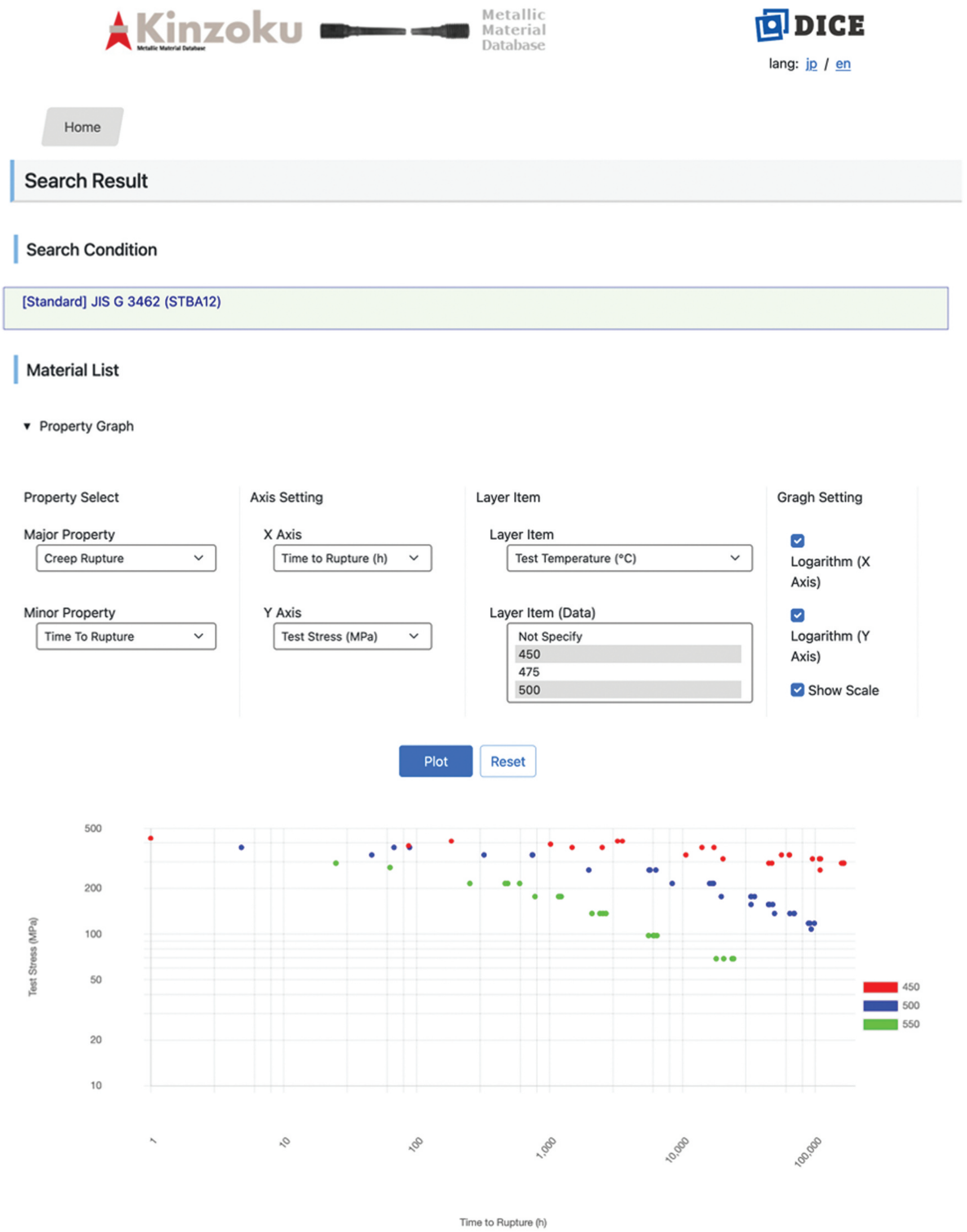


Figure 7. New version of NIMS metallic materials database 'Kinzoku' on MatNavi, using the relational data model described in this paper.

Another point in contrast to RDBMS is that these stores can be extremely flexible in their structures, allowing for different fields for each record when needed. While this is useful in recording various kinds of data, certain level and consistency across the records is also beneficial, which can be achieved by defining a schema. The schema can stipulate the list of keys that are used in the record and the parent–child relationships between the entity types, and the XML records can be validated against this schema. In our schema, the root element is <data> for all records and the first-level child elements are <material>, <process>, <structure>, <test>, and <property> (Figure 8). The <material> element's child elements include those that correspond to the Material table in our spreadsheet and RDB implementations, such as <symbol_of_grade> and <nominal_composition>. The <piece> element is included as a child of <material> to mirror our conceptual data model, a prime example of the self-explanatory nature of this format. Material property data such as creep and fatigue properties are described as values within the generic item element <i>, with specific property names such as '0.2%proof_stress' or 'tensile_strength' given as element attributes. The schemas were developed alongside an extension to Ashino's Materials Ontology [44,45]. Metadata of the XML records were described using Resource Description Framework (RDF), mapped to the Materials Ontology written in Web Ontology Language (OWL) [46]. The Materials Ontology OWL files are available online [47]. Through this arrangement, our data model is connected to a global ontology, which is one of the keys to facilitate materials data integration with various heterogeneous data [7]. An application that can load the OWL/RDF, extract the data structure, and display data from the XML records using XML Path Language (XPath) was developed. More information for this implementation has been reported earlier [48].

5. Conclusion

We have detailed the development of the data model and data format variations for unified data storage and analysis of metallic materials reliability data. Through examining the structures of *NIMS Creep*

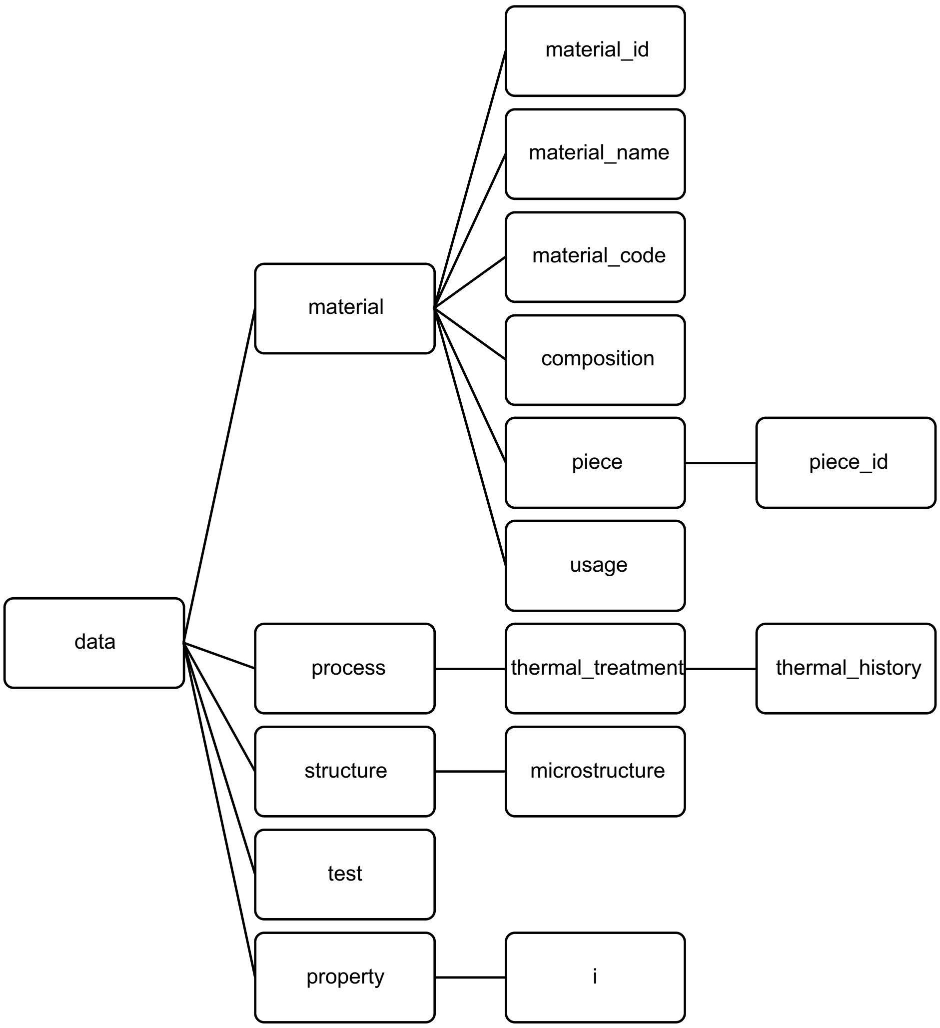


Figure 8. Element hierarchy in the XML representation.

Data Sheet and *NIMS Fatigue Data Sheet*, we designed the model to be capable of addressing various reliability properties and complexities such as welded specimens or heat treatments. Multiple data formats exist for this data model – spreadsheets, RDBMS, and key-value document stores – each offering distinct benefits for robust data management and utilization. Spreadsheets offer low-barrier daily data management and version control using common software tools. RDBMS offers data integrity and high-performance querying capabilities, suitable for system-oriented data management and as a backend to database applications – the prime example being NIMS's renewed Kinzoku database service. Document-oriented key-value stores provide flexible data representation and integration with broader data ontologies. The data model has been shown to be applicable to data from various structural materials reliability initiatives. It was also shown to be capable of supporting cross-organizational data interoperability for a federated data-driven methodology.

Acknowledgements

This work was financially supported in part by the Council for Science, Technology and Innovation (CSTI) Cross-Ministerial Strategic Innovation Promotion Program (SIP) 'Materials Integration for Revolutionary Design System of Structural Materials' (Funding agency: JST) and by Ministry of Education, Culture, Sports, Science and Technology (MEXT) Data Creation and Utilization-Type Materials Research and Development Project (DxMT) JPMXP1122684766. This study was conducted under the National Institute for Materials (NIMS) Structural Materials DX-MOP framework, and by using the Creep Data Sheet (CDS), the Fatigue Data Sheet (FDS), and the Metallic Materials Database (Kinzoku) by NIMS. A.M. wishes to thank Nobumasa Morito for data format discussions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Council for Science, Technology and Innovation; Ministry of Education, Culture, Sports, Science and Technology [JPMXP1122684766]; National Institute for Materials Science; Cross-ministerial Strategic Innovation Promotion Program (SIP), "Structural Materials for Innovation" and "Materials Integration for Revolutionary Design System of Structural Materials".

ORCID

Asahiko Matsuda http://orcid.org/0000-0001-5989-027X
Masahiko Demura http://orcid.org/0000-0002-7308-3041
Takuya Kadohira http://orcid.org/0000-0003-0569-1309
Toshihiro Ashino http://orcid.org/0009-0002-9592-8516
Yoshiyuki Furuya http://orcid.org/0000-0002-3039-5280
Kota Sawada http://orcid.org/0000-0001-7780-1648

References

[1] Himanen L, Geurts A, Foster AS, et al. Data-driven materials science: status, challenges, and perspectives. Adv Sci. 2019;6(21):1900808. doi: 10.1002/advs.201900808

[2] Demura M. Data kudō zairyō kenkyū no shinten – NIMS ni okeru jissen o rei to shite [Data-driven materials research and materials data platforms – an example from NIMS]. J Inf Sci Technol Assoc. 2021;71(6):252–257. Japanese. doi: 10.18919/jkg.71.6_252

[3] Nasiri S, Khosravani MR. Applications of data-driven approaches in prediction of fatigue and fracture. Mater Today Commun. 2022;33:104437. doi: 10.1016/j.mtcomm.2022.104437

[4] Geng X, Wang F, Wu H-H, et al. Data-driven and artificial intelligence accelerated steel material research and intelligent manufacturing technology. Mater Genome Eng Adv. 2023;1(1):e10. doi: 10.1002/mgea.10

[5] Hill J, Mannodi-Kanakkithodi A, Ramprasad R, et al. Materials data infrastructure and materials informatics. In: Shin D Saal J, editors. Computational materials system design. Cham (Switzerland): Springer; 2018. p. 193–225.

[6] Demura M, Koseki T. SIP-Materials integration projects. Mater Trans. 2020;61(11):2041–2046. doi: 10.2320/matertrans.MT-MA2020003

[7] Stier SP, Kreisbeck C, Ihssen H, et al. Materials acceleration platforms (MAPs): accelerating materials research and development to meet urgent societal challenges. Adv Mater. 2024;36(45):2407791. doi: 10.1002/adma.202407791

[8] Demura M. Challenges in materials integration. ISIJ Int. 2024;64(3):503–512. doi: 10.2355/isijinternational.ISIJINT-2023-399

[9] Smith GV. An evaluation of the elevated temperature tensile and creep-rupture properties of wrought carbon steel. Philadelphia (PA): American Society for Testing and Materials; 1970. (ASTM Data Series; DS 11S1).

[10] Smith GV. Supplemental report on the elevated-temperature properties of chromium-molybdenum steels (an evaluation of 2¼Cr-1Mo steel). Philadelphia (PA): American Society for Testing and Materials; 1971. (ASTM Data Series; DS 6S2).

[11] Smith GV. Evaluations of the elevated temperature tensile and creep-rupture properties of C-Mo, Mn-Mo and Mn-Mo-Ni steels. Philadelphia (PA): American Society for Testing and Materials; 1971. (ASTM Data Series; DS 47).

[12] Creep Committee, the Iron and Steel Institute of Japan, editor. Kinzoku zairyō kōon kyōdo data shū dai 1 pen tei gōkin hen [report on the Mechanical properties of metals at elevated temperatures Vol. I low alloy steels]. Tokyo: The Iron and Steel Institute of Japan; 1972. Japanese.

[13] Creep Committee, the Iron and Steel Institute of Japan, editor. Report on the Mechanical properties of metals at elevated temperatures Vol. II stainless

steels. Tokyo: The Iron and Steel Institute of Japan; 1975.
[14] Creep Committee, the Iron and Steel Institute of Japan, editor. Report on the Mechanical properties of metals at elevated temperatures Vol. III carbon steels and cast irons. Tokyo: The Iron and Steel Institute of Japan; 1977.
[15] Creep Committee, the Iron and Steel Institute of Japan, editor. Report on the Mechanical properties of metals at elevated temperatures Vol. IV superalloys. Tokyo: The Iron and Steel Institute of Japan; 1979.
[16] Creep Committee, the Iron and Steel Institute of Japan, editor. Report on the Mechanical properties of metals at elevated temperatures Vol. V deposited metals. Weld metals and welded joints. Tokyo: The Iron and Steel Institute of Japan; 1985.
[17] Holdsworth S. The European Creep Collaborative Committee (ECCC) approach to creep data assessment. J Press Vessel Technol. 2008;130 (2):024001. doi: 10.1115/1.2894296
[18] Sawada K, Kimura K, Abe F, et al. Catalog of NIMS creep data sheets. Sci Technol Adv Mater. 2019;20 (1):1131–1149. doi: 10.1080/14686996.2019.1697616
[19] Furuya Y, Nishikawa H, Hirukawa H, et al. Catalogue of NIMS fatigue data sheets. Sci Technol Adv Mater. 2019;20(1):1055–1072. doi: 10.1080/14686996.2019.1680574
[20] National Institute for Materials Science. Creep Data Sheet (CDS). MatNavi. [cited 2025 Apr 18]. Available from: https://cds.nims.go.jp/
[21] National Institute for Materials Science. Fatigue Data Sheet (FDS). MatNavi. [cited 2025 Apr 18]. Available from: https://fds.nims.go.jp.//
[22] National Institute for Materials Science. Corrosion Data Sheet (CoDS). MatNavi. [cited 2025 Apr 18]. Available from: https://cods.nims.go.jp/
[23] National Institute for Materials Science. Space Use Materials Strength Data Sheet (SDS). MatNavi. [cited 2025 Apr 18]. Available from: https://sds.nims.go.jp/
[24] Silberschatz A, Korth HF, Sudarshan S. Database system concepts. 7th ed. (NY): McGraw-Hill; 2019.
[25] Demura M. Material data kōzō no sekkei: Kanri to katsuyō no balance o ika ni toru ka? [Designing the structures of materials data: how to maintain the balance between management and utilization?]. J Soc Inorg Mater Jpn. 2023;30(427):304–309. Japanese. doi: 10.48505/nims.4744
[26] Demura M, Kadohira T, Ashino T. Chapter 5, Section 1, Kōzō zairyō ni okeru data kōzō no kōchiku to sono ōyō [Construction of data structure for structural materials and its application]. In: Demura M, Enoki M Watanabe M, et al., editors. Materials integration ni yoru kōzō zairyō sekkei handbook [Materials integration for structural materials design]. Tokyo: NTS; 2024. p. 263–282. Japanese.
[27] General comments on the program and procedure for creep and rupture tests on high-temperature materials manufactured in Japan. Tokyo: National Research Institute for Metals; 1972. (NRIM Creep Data Sheet; 0). Available from: https://cds.nims.go.jp./
[28] Yamazaki M, Xu Y, Fujita M, et al. Internet o katsuyō shita zairyō jōhō no teikyō oyobi kensaku system no kōchiku [Development of internet materials information and search system]. Joho Chishiki Gakkaishi. 2005;15(3):11–18. Japanese. doi: 10.2964/jsik_KJ00003803573
[29] Yamazaki M, Xu Y, Murata M, et al. NIMS structural materials databases and cross search engine – MatNavi. In: Veivo J Auerkari P, editors. BALTICA VII – life management and maintenance for power plants. Espoo: VTT Technical Research Centre of Finland; 2007 [cited 2025 Apr 18]. p. 193–207. (VTT Symposium 247). Available from: https://publications.vtt.fi/pdf/symposiums/2007/S247.pdf
[30] Yagi K. Acquisition of long-term creep data and knowledge for new applications. Int J Press Vessels Pip. 2008;85(1):22–29. doi: 10.1016/j.ijpvp.2007.06.001
[31] Data sheets on the elevated-temperature properties for base metals, weld metals and welded joints of 18Ct-8Ni stainless steel plates (SUS 304-HP). Tokyo: National Research Institute for Metals; 1995. (NRIM Creep Data Sheet; 32A). NRIM Creep Data Sheet; 32A https://cds.nims.go.jp/
[32] Data sheets on the elevated-temperature properties for base metals, weld metals and welded joints of 18Cr-12Ni-Mo-middle N-low C hot rolled stainless steel plate (SUS 316-HP). Tsukuba: National Institute for Materials Science; 2005. (NIMS Creep Data Sheet; 45A). NIMS Creep Data Sheet; 45A https://cds.nims.go.jp/
[33] Program of the fatigue data sheet project for engineering materials manufactured in Japan. Tokyo: National Research Institute for Metals; 1978. (NRIM Fatigue Data Sheet; 0). Available from: https://fds.nims.go.jp/
[34] Data sheets on fatigue crack propagation properties for butt welded joints of SB42 carbon steel plate for boilers and other pressure vessels –effect of stress ratio–. Tokyo: National Research Institute for Metals; 1984. (NRIM Fatigue Data Sheet; 41). Available from: https://fds.nims.go.jp/
[35] Nagahashi K, Shimura K, Otoguro Y. Atsuryoku yōki-yō zairyō (Cr-Mo kei kō) database system [A database system for pressure vessel materials (Cr-Mo steels)]. J Inf Process Manag. 1997;40(1):28–37. Japanese. doi: 10.1241/johokanri.40.28
[36] Nagahashi K, Iijima K, Matsubara M, et al. Atsuryoku yōki zairyō database system no kaihatsu [Development of database system for pressure vessel materials]. Therm Nucl Power. 1998;49(1):92–102. Japanese.
[37] Zhang Z, Tang H, Xu Z. Fatigue database of complex metallic alloys. Sci Data. 2023;10(1):447. doi: 10.1038/s41597-023-02354-1
[38] Sakurai J, Demura M, Inoue J, et al. Kikai gakushū ni yoru ferrite kei tainetsukō no creep hadan jumyō yosoku [Creep life predictions by machine learning methods for ferritic heat resistant steels]. Tetsu-To-hagané. 2022;108(7):424–437. Japanese. doi: 10.2355/tetsutohagane.TETSU-2022-003
[39] Sakurai J, Demura M, Inoue J, et al. Creep life predictions by machine learning methods for ferritic heat resistant steels. ISIJ Int. 2023;63(10):1786–1797. doi: 10.2355/isijinternational.ISIJINT-2023-266
[40] Sakurai J, Torigata K, Matsunaga M, et al. Creep hadan jikan oyobi kōon hippari kyōdo yosoku model no rengō gakushū [Federated learning of creep rupture time and high temperature tensile Strength prediction models]. Tetsu-To-Hagané. 2025;111(5):246–262. Japanese. doi: 10.2355/tetsutohagane.TETSU-2024-124

[41] National Institute for Materials Science. Metallic Material Database (Kinzoku). MatNavi. [cited 2025 Apr 18]. Available from: https://metallicmaterials.nims.go.jp/

[42] Ogata T, Yamazaki M. New stage of MatNavi, materials database at NIMS. In: Zhao J-C, Asta M, Gumbsch P, et al. editors. Harnessing the materials genome: accelerated materials development via computational and experimental tools. Vail (CO): Engineering Conferences International. ECI Symposium Series; 2013. p. 146–161. [cited 2025 Apr 18]. Available from: https://dc.engconfintl.org/materials_genome/9

[43] Yamazaki M, Hosoya J, Ogata T, et al. NIMS metallic materials database: Kinzoku. In: The 4th Asian Materials Data Symposium; Jeju, Korea; 2014.

[44] Ashino T, Fujita M. Definition of a web ontology for design-oriented material selection. Data Sci J. 2006;5:52–63. doi: 10.2481/dsj.5.52

[45] Ashino T. Materials Ontology: an infrastructure for exchanging materials information and knowledge. Data Sci J. 2010;9:54–61. doi: 10.2481/dsj.008-041

[46] OWL Working Group. Web Ontology Language (OWL). Semantic Web Standards. 2012 [cited 2025 Apr 18]. Available from: https://www.w3.org/OWL/

[47] Ashino T. MaterialOntology. 2014 [cited 2025 May 14]. Available from: https://www.researchgate.net/publication/262004148_MaterialOntology

[48] Ashino T, Nishikawa N, Kadohira T, et al. Implementation of materials data integration using ontology. In: Pozanenko A, Stupnikov S Thalheim B, et al. editors. Supplementary proceedings of the XXIII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2021); Moscow (Russia). CEUR Workshop Proceedings; 2021. p. 235–239. [cited 2025 Apr 18]. Available from: https://ceur-ws.org/Vol-3036/paper18.pdf