

Material Database for Efficient Development Using Materials Informatics

Tsuyoshi ITO¹
Nao SUGIMOTO¹

Satoko ISHII¹
Yoichi TOYOOKA¹

Hiroshi HONDA¹
Toru FURUSAWA¹

ABSTRACT

A comprehensive material database that enables the storage, management and export of material data and related information in a standardized format has been constructed to promote the use of materials informatics to achieve highly efficient materials research, such as the application of information science to predict physical properties in order to reduce experimental trial and error.

Usability and user differences such as expertise were taken into account when selecting the database tools and creating the internal structure. As a result, storage of around 10,000 test data and standard properties in various material fields has been completed.

It was confirmed that this enables a reduction in the work time required for data cleansing, which is an issue for use of materials informatics. It also facilitates the acquisition of data needed for design and CAE, etc., and helps to increase operations efficiency and reduce development costs.

1. Introduction

Due to factors such as economic globalization and the rapid progression of global warming, there are growing needs for the enhancement of automotive parts, including weight and cost reduction and higher functionality. Material and manufacturing process development has also been promoted for the materials that comprise automotive parts in order to achieve replacement with aluminum and resin materials that have lower density than steel materials⁽¹⁾⁻⁽³⁾. This type of development process typically adopts the empirical method, in which various prototype materials with different composition and manufacturing conditions are synthesized, and the basic properties are obtained using test pieces. However, functional materials such as adhesives, catalysts, and magnets, and new manufacturing processes such as additive manufacturing (AM)⁽⁴⁾ have a wide range of factors that influence material properties, and this entails enormous cost and work time to acquire the basic properties for analysis of those factors. Therefore, methods that use materials informatics (MI), which integrates materials science and machine learning to predict material properties from test results, are attracting attention⁽⁵⁾.

MI uses material properties as target variables and factors such as composition and manufacturing conditions as explanatory variables, performs statistical processing to derive the relationships between these variables, and

optimizes the material composition and manufacturing process based on those relationships. Therefore, prerequisites for use of MI include not only a certain amount of material property measurement data, but also the necessary related information and an integrated data format. In practice, however, there are individual differences in the expertise, roles, and skills, etc., of data owners, so information such as data distribution, test conditions, and units may be missing, and there may be variations in notation, such as the use of trade names or abbreviations as material names. Therefore, in cases such as when there are multiple data owners, first the owner is identified, and then the data is aggregated, data with the necessary related information is selected, and a data list is created with an integrated format such as data arrangement, unit, and notation of terms. Further processing, such as data conversion according to the purpose, is then performed, enabling the start of data analysis. This type of data cleansing work to collect, convert, and organize data is said to account for 70% of the work time needed for machine learning, including MI, and is a major issue for the use of MI^{(6),(7)}.

2. Development Goals

Efficient material development using MI calls for swifter data cleansing by storing material test data structurally, together with manufacturing conditions and other incidental

¹ Honda Motor Co., Ltd.

Table 1 Comparison table of DB software

	Requirements	DB Software		
		Software [A]	Software [B]	GRANTA MI
Input	Material name, standard name and product name can be input together	+++	+++	+++
	Attributes can be added easily, and input template is created automatically	+	++	+++
	Unit can be converted automatically	+++	+++	+++
	Links between data can be created easily	+	++	+++
Display	Display items and layout can be changed according to the purpose	+	+++	+++
	Graphs can be superimposed easily	+	+++	+++
	Graphs can be created from the results of search and statistical processing	+	++	+++
Search	Complex search including process and test conditions can be performed	+	++	+++
Export	Data can be exported in Excel, PDF, CSV format, etc.	+	+++	+++
	Works with Abaqus, LS-Dyna, FEMFAT, Nastran, Enovia, CATIA, etc.	++	+++	++
Administration	Access can be controlled by data, folder or database	+	+++	+++
	Revision and release history remains	+	+++	+++
	Easy modifications for large amounts of data	+	+	+++
	Reliability rank of data can be displayed	+	+++	+++
Cost	Reasonable purchase price and maintenance cost	+++	+	+++

information in a material database (DB) and exporting these data in a fixed format. Furthermore, if representative material properties used for design and CAE can be centrally stored and shared in a material DB together with material test data, various operational efficiency enhancements such as standardization of data, reduction of inquiry work time, sharing of know-how, and leveling of workload can be expected in addition to reduction of development costs.

Therefore, the authors decided to construct a new material DB with the following two goals.

- (1) Swifter data cleansing for efficient MI use.
- (2) Centralized management of material data for enhancement of operational efficiency and reduction of development costs.

3. Material DB for Use of MI

When using material DB that are publicly accessible such as via the Internet in this development, it is a challenge to store in-house data in a flexible manner that maintains confidentiality^{(8), (9)}. In addition, many material DBs operated within companies list specific material properties such as Young's modulus and density, and storage of other material properties and detailed test data would entail large system renovation costs. Furthermore, in order to construct a material DB and maintain long-term functionality, it is necessary to not only store the information sought by users, but also to take into account convenience by enabling easy data registration, modification, search, and utilization in accordance with changing needs.

Therefore, commercially available DB software that enables DB construction without specialized skills such as programming languages and that also enables flexible

modification was used to design a highly convenient DB structure, and data was accumulated.

3.1. Selection of DB Software

When selecting the DB software, multiple software with proven records both in-house and outside the company were evaluated beforehand in terms of input, display, search, export, administrative functions, and usage cost, etc. Table 1 summarizes these results. Software B and ANSYS, Inc.'s GRANTA MI received nearly the same rating, but importance was placed on features such as automatic creation of input templates, search and export with complex conditions, and batch modification of data, and in the end GRANTA MI was selected.

Figure 1 shows an overview of the GRANTA MI system. Like a typical Web application, each user and DB

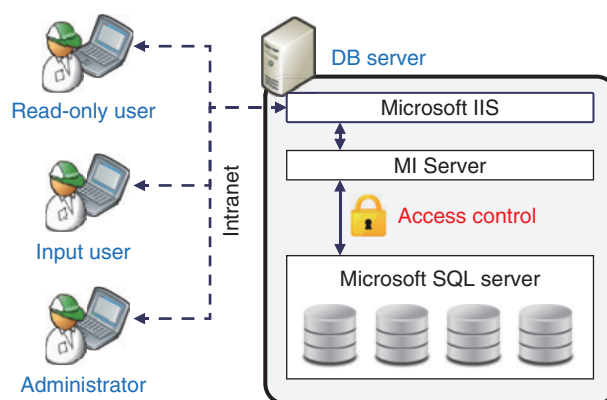


Fig. 1 System configuration of GRANTA MI

administrator logs in to Microsoft IIS from a client PC via an intranet, the GRANTA MI core system MI Server determines the access rights and performs data processing, and multiple DBs within Microsoft SQL Server can then be accessed. Administrators can also limit the data viewing, input and export ranges according to differences in user responsibilities and expertise.

3.2. DB Structure Design

Figure 2 shows various automotive parts and the corresponding required properties for each material type. As shown here, the material properties to be handled differ for each material field, such as strength and rigidity for metals, weather resistance and viscosity for resin materials, and magnetic properties for magnetic materials. This results in a correspondingly diverse range of test methods and registration items. In addition, there are also various manufacturing methods such as casting, rolling, sintering, injection molding, and AM, and each method uses different parameters, so it is more efficient to manage these using separate DBs.

However, representative property values and judgment criteria values for a wide range of materials are used in design work and CAE analysis, but those values are also based on large amounts of test data from different suppliers and test environments, etc. For example, when setting fatigue strength criteria for a material, 10 to 30 tests are conducted at each temperature in the anticipated operating temperature range, and the influence of strength variation due to factors such as surface roughness, hardness, and quality is also taken into account. This calls for separate management of representative values and actual data.

Therefore, the DB structure was separated into a Design and simulation DB that manages representative properties and DBs, such as a Metal DB and Resin DB that manage

the test data for each material field as shown in Fig. 3. The Design and simulation DB structure is further divided into mutually interlinked areas that manage physical property values, together with supplemental information such as representative values and guaranteed values (Design table), manage property values by CAE purpose (Simulation table), and manage applied materials by part (Application table). Meanwhile, it is presumed that material tests are conducted under multiple manufacturing conditions even for the same material type. Therefore, each DB for managing test data was designed with mutually interlinked areas that manage material name lists (Material list table), manage related information such as compositions and manufacturing conditions (Pedigree table), and manage various test conditions and data (Test data table).

This lets design and CAE staff swiftly access the representative material properties needed for operations even if they have little materials expertise, and raw data can also be checked retroactively as needed. It also enables flexible input and export of special test data and manufacturing conditions that differ for each material field, which makes it easier for materials staff to accumulate and reuse data.

3.3. Addition of Data Input Assist Functions

One feature of the above-mentioned GRANTA MI is that the input template is automatically created when the input items are set. However, at the default setting there are no restrictions on the input contents, so when unspecified multiple users freely enter data, differences occur in the material classification method and the degree of data completeness, which does not lead to a reduced data cleansing load. In the case of resin materials, for example, it is necessary to accurately input various classification information such as the base resin, ISO standard, in-house

	Metals	Polymers	Magnetics and catalysts	Chemical
Parts	 Body, cylinder block, gear, etc.	 Bumper, tire, oil, engine cover, etc.	 Rotor, stator, converter, etc.	 Door, color, seat, etc.
Material	Steel, aluminum, magnesium	Resin, rubber, oil, long life coolant	Ceramics, magnet, precious metal	Adhesive, paint, fabric
Demanded properties	Fatigue strength, elongation, rigidity, weldability, heat resistance, etc.	Environmental resistance, wear resistance, fluidity, etc.	Core loss, flux density, conversion efficiency, etc.	Adhesion strength, anti-corrosion performance, viscosity, etc.

Fig. 2 Material classification and required properties

standard, and supplier grade name. In the case of metal materials, it is necessary to keep load conditions such as axial and bending directions together with fatigue test results. However, if this information is missing or incorrect, not only does data correction become necessary, but it can also lead to inconvenience such the inability to find the needed data.

Therefore, input assist functions such as the following were added to the input template to help correctly search for the needed data.

- (1) Item description: Descriptions and actual examples of items are shown in order to minimize differences in interpretation of item names.
- (2) Selection list setting: Items with limited entry contents such as classification are selected from pull-down menus in order to achieve consistent input word notations.
- (3) Image file name input: A function that automatically enters image file names to the template was added in order to streamline the input of large numbers of images.
- (4) Automatic entry: Items such as ID that are to be entered in accordance with administrative rules are entered automatically.
- (5) Contents check: Prior to import, it is possible to check that the data format of items such as text, values, and date are correct, and that required fields are entered.

3.4. Centralized Management of Material Data

When using MI, the more experimental data that can be prepared, the better. However, material data involves the synthesis of prototype materials and creation of test pieces, so it is not an easy task to obtain thousands or

Table 2 Number of records

Material categories	Databases	
	Design and simulation	Metal, resin, etc.
Metals	755	7104
Resins	304	363
Rubbers	87	495
Magnetics	121	785
Fluids	49	66
Catalysts	0	394

tens of thousands of data. Therefore, the Smart Materials Informatics method⁽⁵⁾, which enables the prediction of properties even with a small number of data, was used to investigate the number of data that enables prediction. In the referenced study, artificial data was used to understand the relationship between the number of data and prediction error, and verification was performed using actual catalyst, magnet, and thermoelectric material data. As a result, it was stated that at least several tens to hundreds of data are needed per case for prediction to function effectively⁽⁵⁾. Based on this, the target number of accumulated data in this DB was set as a total of 10,000 records for the 500 main material types used in automobiles, and activities to promote input such as hands-on in-house seminars were conducted. As a result, registration of records equivalent to the target number was completed over the approximately three-year period through April 2020, as shown in Table 2. This is an amount comparable to the 1,130 records of the aerospace material DB (MMPDS) and 6,600 records of the metal material DB (MI-21).

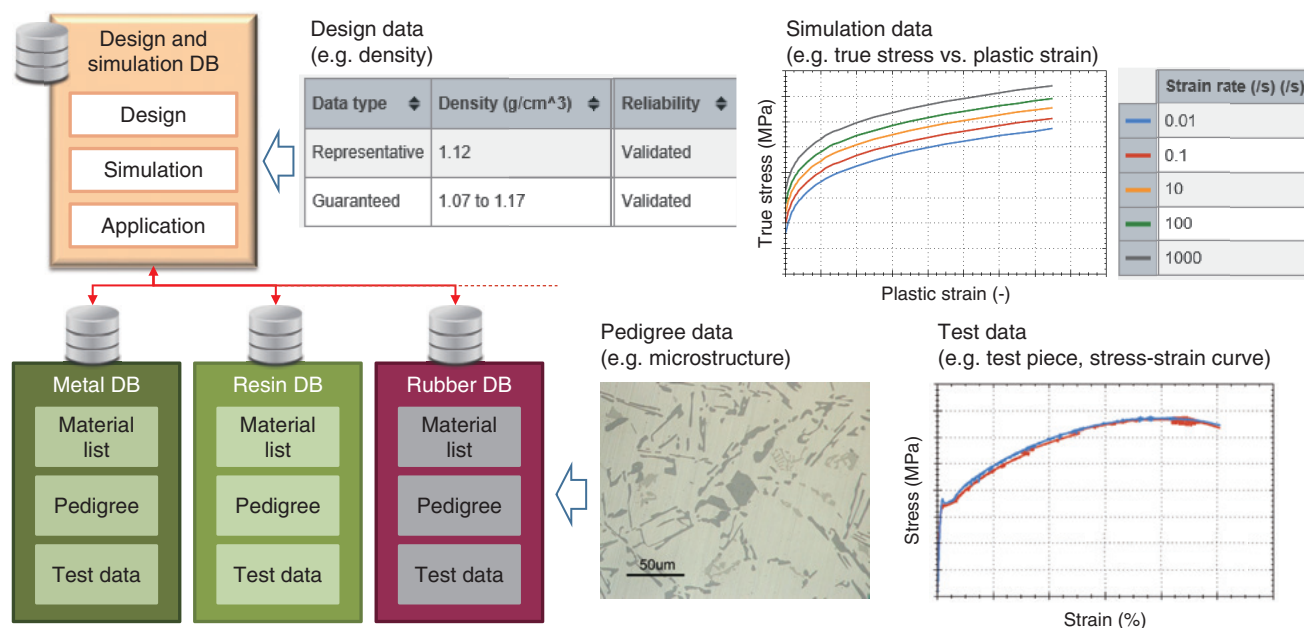


Fig. 3 Structure of database and example data

4. Introduction of Cases of Technical Application

4.1. Use of MI to Promote New Material Development

In order to use MI, it has been necessary to spend much work time on data cleansing, such as collection and selection of test data with clear material manufacturing and test conditions, and integration of data formats. However, the accumulated test data can be stored in this DB and exported in a standardized format together with related information, which enables the use of MI with greatly reduced data cleansing. Figure 4 shows an example of the results for storage and export of chemical material test data. In this example, the measurement results for adhesion strength and viscosity are stored in a linked format with information of the combination of compounds that constitute the chemical materials. In addition, the results of 25 different tests are exported in list format. Although it is still necessary to convert the data as before, the time needed for data collection and selection and format integration was greatly reduced by inputting it in accordance with the input template, which reduced the total data cleansing time by 80% or more as shown in Fig. 5.

Next, Fig. 6 shows the material property prediction results obtained by applying MI to these test results. Here, the correlation between each component and each material property was obtained from 25 test results exported from the DB, and the unknown adhesion strength and viscosity for 400,000 different combinations of components was then predicted based on those correlations and plotted. This figure

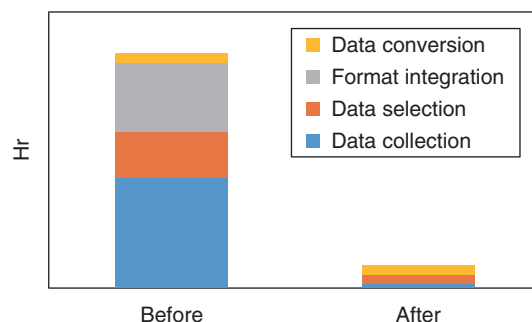


Fig. 5 Change in data cleansing time

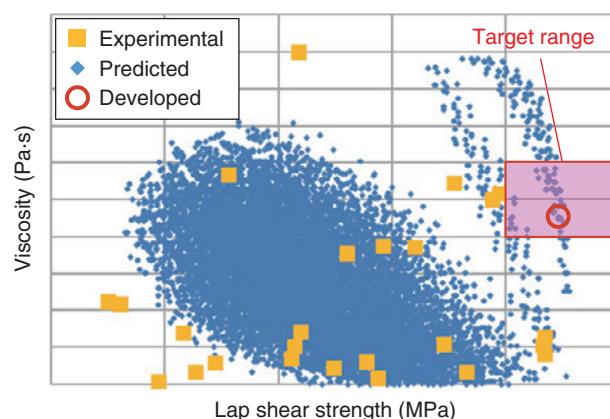


Fig. 6 Property predicted results of adhesive materials

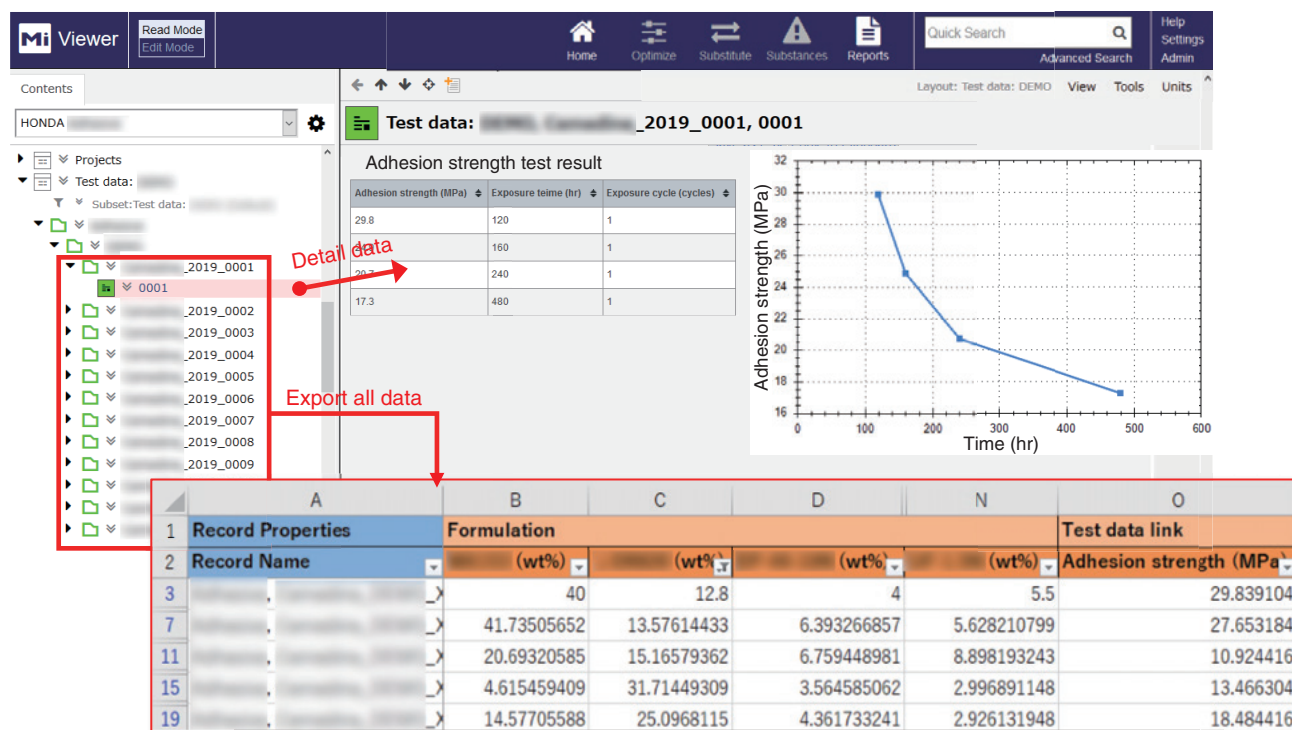


Fig. 4 List output of adhesive test data

shows that combinations of components that can provide high adhesion strength within the viscosity range demanded by automotive production lines can be efficiently extracted. As an example, Fig. 7 compares the measured and predicted values of the material properties of the developed material shown in the red circle in Fig. 6. Thus, the use of data accumulated in this DB together with MI enabled the prediction of unknown material properties with high accuracy, and provided a direction for the development of new materials.

4.2. Enhancement of Operational Efficiency

As part of their duties, materials staff verify functions in consideration of the operating environmental load of each part. For example, the functional life of rubber materials differs according to the operating temperature and the material type, so life verification is performed for each individual part, and alternative materials are investigated from among a large number of material types as necessary. In this case, it was previously necessary to collect heat resistance data for all of the materials to be compared and to compare them by drawing on Excel. However, this DB enables a superimposed display of the heat resistance properties of each material obtained from the accumulated data, as shown in Fig. 8, so materials suited to the operating environment can be investigated from among a wide range of materials in a shorter time. Furthermore, approximate properties can also be estimated from past data even for materials with unknown material properties. Figure 9 shows an example of low- and high-frequency core loss values extracted from 160 separate magnetic property measurement data for electric steel sheets and expressed in an X-Y graph. This figure shows that differences in the correlation of each property according to the sheet thickness can be quantitatively expressed, which enables more accurate estimation even for materials with a thickness for which the properties have not yet been measured.

Use of this DB can be expected to increase design work efficiency when selecting material types based on material properties. Figure 10 shows an example of the results of extracting materials with a yield strength of 80 MPa or more and a tensile strength of 100 MPa or more as representative

properties at 800°C from over 200 types of registered steel material data. This DB also enables a search for the required property value range in this manner, and also a search for materials by incidental information such as guaranteed values and representative values, which lets designers narrow down the material candidates without inquiries to specialist materials staff. In addition to increasing the work efficiency of designers, this is expected to promote the in-house spread of design studies with standardized physical property values, leading to higher design quality and reduced development reworking.

This DB also stores material models by CAE purpose, so for example, material properties for crash analysis used in domestic automobile development can be used for the same application at overseas development sites. This is expected to increase CAE work efficiency, and also enables a reduction in the cost of material testing, which has been duplicated by each department in order to construct material models. As a result, even adding in license, server administration and other costs related to material DB introduction, it is estimated that material property acquisition costs can be reduced by approximately 41% as shown in Fig. 11.

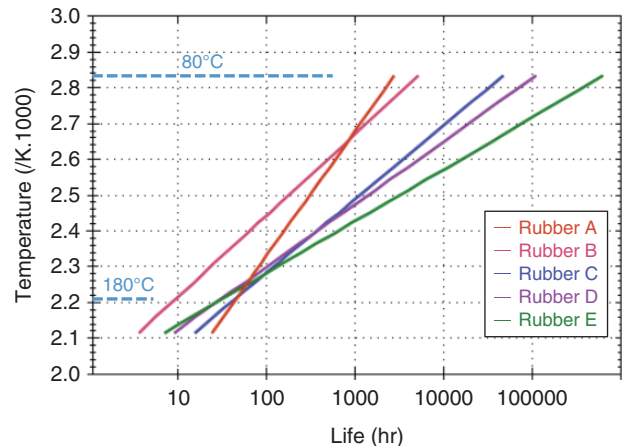


Fig. 8 Temperature and lifetime of rubber materials

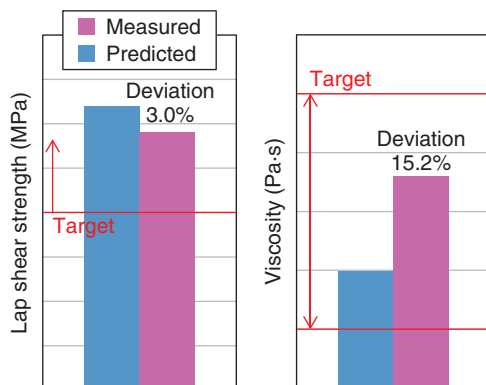


Fig. 7 Properties of developed material

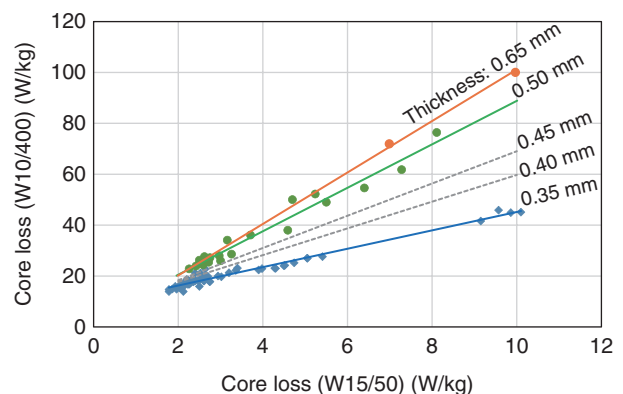


Fig. 9 Relationship of core losses by frequency of electrical steel sheet

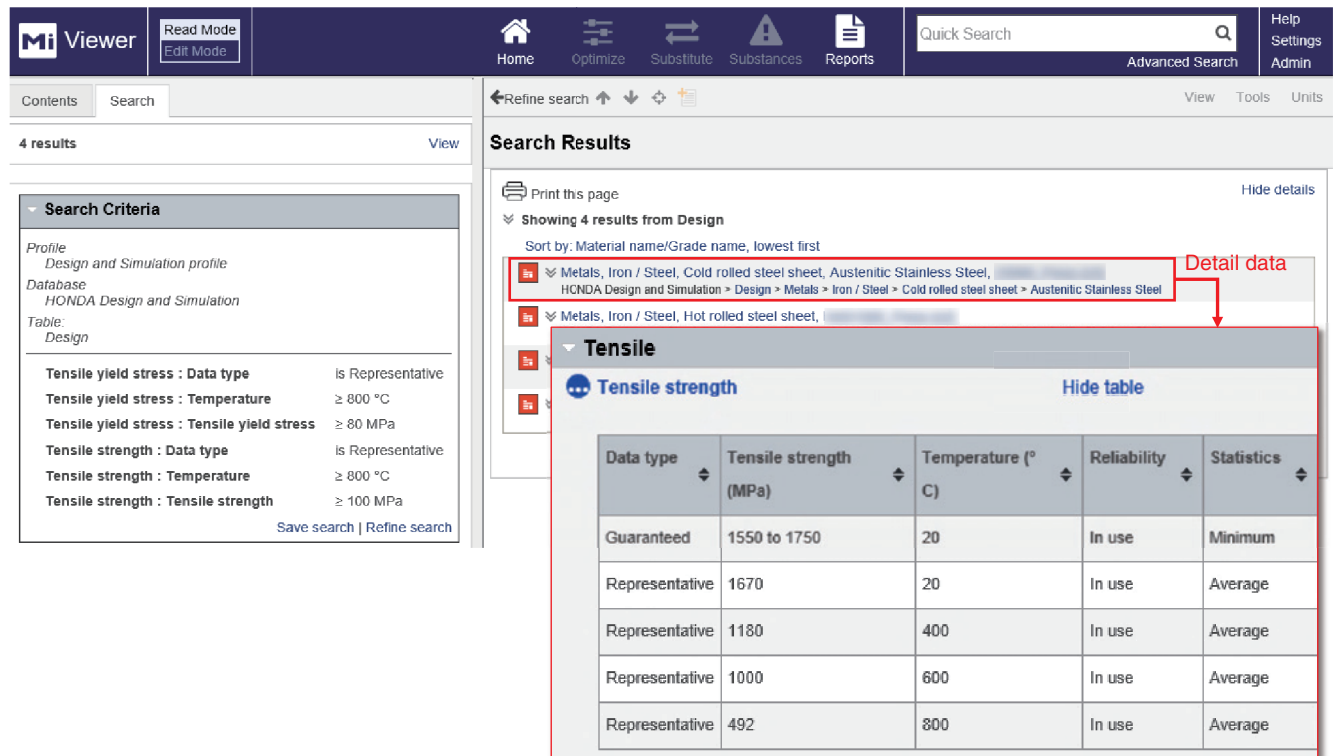


Fig. 10 Search results by material properties

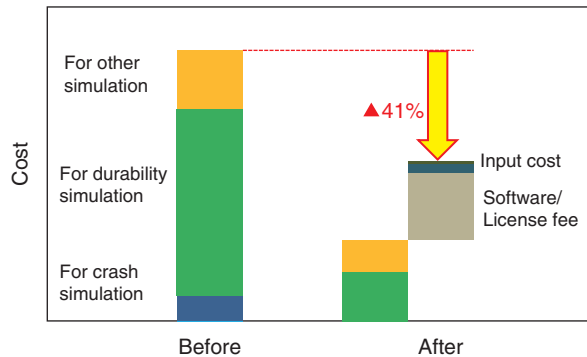


Fig. 11 Material property measurement cost for simulation model

5. Conclusion

A material DB that can store, manage, and export various in-house material data structurally, together with incidental information, has been designed and constructed to increase the efficiency of development operations by utilizing MI. The following effects were obtained as a result.

- (1) The use of data accumulated in this DB together with MI enabled the prediction of unknown material properties with high accuracy, and provided a direction

for the development of new materials.

- (2) By storing and exporting the test data for each material field structurally, together with related information, the work time for data cleansing was greatly reduced, which was an issue when using MI.
- (3) Use of a structure that manages the standard material properties used for design and CAE separately from material test data enables each staff to access the needed material properties by themselves. In addition, it was shown that the cost of acquiring material properties for the construction of material models can be reduced by sharing data between domestic and overseas development departments.

References

- (1) Imanishi, T., Asami, A., Saneyoshi, Y.: Large-sized and Lightweight Aluminum Hollow Diecast Subframe Using High Pressure Diecasting, Honda R&D Technical Review, Vol. 29, No. 1, p. 46-51
- (2) Yamane, Y., Sakuraba, T.: Development of Lightweight CFRP Hood and Rear Spoiler, Honda R&D Technical Review, Vol. 15, No. 1, p. 209-214
- (3) Washizu, K., Takeishi, I., Sato, N., Ogishi, H., Tabata, M.: Development of High-pressure Diecast Magnesium Intake Manifold, Honda R&D Technical Review, Vol. 17, No. 2, p. 32-37
- (4) Asanuma, M., Endo, T., Sakai, H.: Introduction of Heat

Exchanger Production Technique for Stirling Engine Using Additive Manufacturing, Honda R&D Technical Review, Vol. 28, No. 1, p. 174-182

- (5) Furukawa, A., Ikeda, T., Okayama, T.: Materials Research Method using Smart Materials Informatics, Honda R&D Technical Review, Vol. 29, No. 1, p. 90-97
- (6) <https://findy-code.io/engineer-lab/machine-learning-tough-work>
- (7) Norio, S.: Education and Human Resource Development of Machine Learning in Toyota Motor Kyushu, Journal of the Japanese society for quality control, Vol. 49, No. 2, p. 160-164, (2019) (In Japanese)
- (8) Metallic Materials Properties Development and Standardization (MMPDS), 14th Edition, Federal Aviation Administration, p. 2535, (2019)
- (9) <https://www.totalmateria.com/>

■ Author ■



Tsuyoshi ITO



Satoko ISHII



Hiroshi HONDA



Nao SUGIMOTO



Yoichi TOYOOKA



Toru FURUSAWA