"Trustworthy and Explainable AI" Achieved Through Knowledge Graphs and Social Implementation

● Masaru Fuji ● Katsuhito Nakazawa ● Hiroaki Yoshida

While AI performance continues to improve, when experts use it in practical applications, its presented inferences may not provide sufficient information for decision-making. This failure to ensure sufficient reliability has encouraged R&D on explainable AI to be conducted on a global scale in recent years. In response to this problem, Fujitsu has achieved the world's first "trustworthy and explainable AI" that makes use of knowledge graphs capable of representing expertise in a systematic manner to explain the reason and basis. This provides users with sufficient information to make decisions and improves AI reliability. This paper outlines knowledge graphs as a basis for Fujitsu's trustworthy and explainable AI. It also introduces application examples in the fields of finance and chemistry.

1. Introduction

When applying AI in an industrial field that requires a high degree of reliability such as medical treatment, finance, or chemistry, the answer derived from AI will be put to use after being interpreted by an expert in that field. However, for an expert using AI, an AI inference simply in the form of "yes/no" or "90%" does not in itself provide sufficient information for making a decision. As a result, the inability to make a reliable decision as to whether to use an AI inference has become a problem. In response to this issue, research of "explainable AI" has become quite active in recent years.

To date, the field of explainable AI has revolved around the efforts to improve explainability by adding explanatory functions to machine learning centered about deep learning. These efforts on their own, however, have yet to sufficiently guarantee the degree of AI reliability needed in industrial fields.

Taking this into account, Fujitsu has become the first in the world to achieve the social implementation of "trustworthy and explainable AI" that improves reliability by providing a reason or basis for an AI inference through the use of knowledge graphs.

In this paper, we first provide an overview of knowledge graphs as a basis for trustworthy and explainable AI and then introduce application examples in the fields of finance and chemistry.

2. Explainable AI and related issues

This section describes the concept of explainable AI and discusses related issues.

2.1 What is explainable AI?

AI can learn statistical features from data and make predictions of unknown phenomena. Here, AI that can explain the reason or basis behind its thinking in arriving at a prediction is called "explainable AI."

Current research and development of explainable AI is focused on deep learning using AI learning and inference and is mainly concerned with identifying which statistical features in input data become deciding factors in an AI inference. For example, to represent AI thinking in determining whether a certain photo includes the image of a cat, the method used by one type of explainable AI called "deep explanation" is to present the activation level of an AI neuron that recognizes the ear portion of a cat as a feature. An AI neuron, which corresponds to a nerve cell in the human brain, is one of many nodes that make up a neural network used in deep learning.

In this example involving image recognition, the ear portion of a cat was identified as being the deciding factor, so if that can be presented to the user superposed on the input image, the user can visually surmise the reasoning behind that AI decision. This direction of research is making rapid progress. For example, in fields in which it is easy for humans to make decisions on a visual basis as in recognizing an object in an image, this research is producing results that are approaching a practical level.

2.2 Related issues

As described above, presenting the reasoning behind an AI decision improves explainability by a certain amount. However, the reality is that this by itself cannot guarantee the level of reliability required at actual worksites in industry and elsewhere. To guarantee reliability, it is desirable to present the AI-based inference together with its basis that an expert in that field can reference.

In the field of medical treatment, for example, let’s assume a system in which AI is used to infer a diagnosis when using large amounts of patient data and associated diagnoses as training data. In this case, it would be insufficient to simply present to the physician which patient data provided the reason for the AI decision. It is important here to present that information together with references to journal papers or other material that provide a deep basis for that decision to enable an expert to make better use of that AI inference.

3. Trustworthy and explainable AI

Fujitsu Laboratories has achieved the world’s first trustworthy and explainable AI that improves AI reliability by giving the basis for an AI inference using knowledge graphs. Here, knowledge graphs can be constructed using academic literature and documents as sources of expert knowledge.

The configuration of explainable AI based on knowledge graphs is shown in **Figure 1**. In this configuration, Deep Tensor inputs graph-structured data shown in the upper-left portion of the figure as training data and outputs an inference result and inference factors (reason). In addition, the knowledge graph shown in the lower half of the figure gives the basis for this result.[1]

Knowledge must be made into a computer-readable form to provide a reason or basis for an AI inference on the basis of expert knowledge as found in academic literature and other documents. In the research presented here, we represented knowledge on a computer in the form of knowledge graphs, which is a framework for systematically representing knowledge in a specialized field.

There are a variety of techniques for representing knowledge, and one of these is a method that represents knowledge in the form of connections between

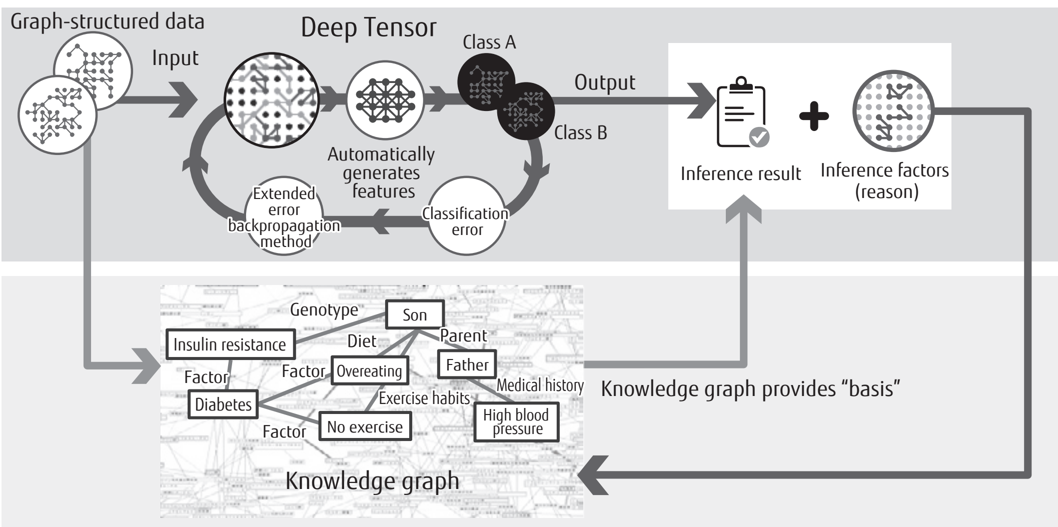


Figure 1
Configuration of explainable AI based on a knowledge graph.

concepts. The knowledge graph is a good example of this method. A method that represents knowledge as connections between concepts tends to have a simple structure, which makes it conducive to mechanical processing as in AI.

The configuration of a knowledge graph is shown in **Figure 2**. In the lower half of the figure, a knowledge system is expressed as connections between concepts, and in the upper half, field-unique definitions related to associations between these concepts are indicated in an ontological form that provides explicit specifications for conceptualization. For example, knowledge that overeating is one cause of diabetes can be expressed as the connection "cause" between the two concepts of "overeating" and "diabetes." Adding meaning to a knowledge graph through the use of metadata, ontology, etc. enables AI to provide a deeper understanding of connections between concepts. Here, moving the knowledge possessed by mankind to knowledge graphs can configure knowledge in a form that incorporates the knowledge of many individuals beyond the knowledge that any one human being could handle.

In addition, field-unique definitions in the upper half of Figure 2 are those described by experts in that field. Moreover, connections between concepts in the lower half consist of those created from openly released "public data" such as academic literature and web-based information and those created from unreleased "non-public data" that are held inside an organization such as a company or university. In the figure, Linked Open Data (LOD) refers to a method of releasing and sharing public data on the Internet.

4. Application example in finance

Fujitsu Laboratories is promoting the application of its recently developed trustworthy and explainable AI to actual fields in industry requiring a high level of reliability such as genomic medicine, finance, and chemistry. This section describes an application example in the field of finance. An application example in the field of genomic medicine[2] is described in a separate paper.[3]

In the field of finance, recent years have seen increasing activity in financial technology (FinTech) that aims to provide financial products and services that make use of ICT. For example, studies are looking at ways of automating the screening of loan applications at banks by introducing AI into credit scoring with the aim of making loan operations more efficient.[4]

On the other hand, introducing AI into a mission-critical field such as finance requires that a result output from AI be accompanied by a highly convincing

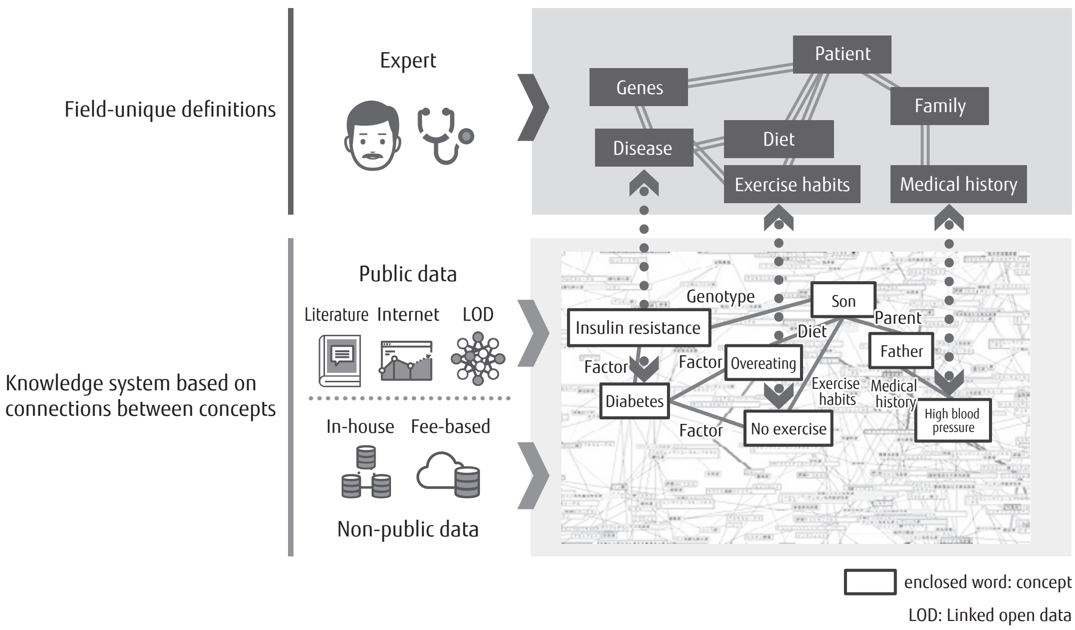


Figure 2
Configuration of a knowledge graph.

explanation. The inability to completely solve this problem has been a major obstacle to the introduction of AI in this field.

At present, a common approach to AI-based loan-application screening is to give the reason for the result output by AI by a method that identifies the features that affect that result. However, in the event that the borrower asks for an even more convincing explanation from the bank, there is a need for deeply trustworthy and explainable AI that can give the basis for judging those features to be important at a level that people can understand.

At Fujitsu Laboratories, we developed technology for identifying a growing company from its credit score with the aim of applying trustworthy and explainable AI to the field of finance. This technology inputs exchange-transaction data between companies possessed by a bank after converting it to a graph structure and outputs inference factors from Deep Tensor. Using these inference factors and a financial knowledge graph, we can explain the basis for arriving at the output result.

A financial knowledge graph is a knowledge base in graphical form that is configured by comparing information that corporate numbers publically released by the National Tax Agency JAPAN and open data released by the Electronic Disclosure for Investors NETwork (EDINET) of the Financial Services Agency in Japan with corporate information possessed by banks.[5),6)] In this way, trustworthy and explainable AI can be provided even in the mission-critical field of finance by not only outputting a highly accurate result of loan screening by AI but also giving a clear basis for the inference factors associated with that result.

A tool for analyzing the basis for judging a loan application from a financial knowledge graph is shown in **Figure 3**. This analysis tool features three main panels of information. The upper-left section of the screen consists of a prediction-contribution list that arranges in descending order the contribution ratio, i.e. degree of influence, of inference factors output by Deep Tensor with respect to the output result. Next, the lower-left section of the screen displays specific inter-company transaction information, i.e. time-series data of transaction volume and number of transactions, between inference factors. Finally, the right section of the screen displays the relationship between inference factors, such as transaction, director, affiliate company, and major shareholder, in graphical network form.

We here introduce the case of analyzing the credit

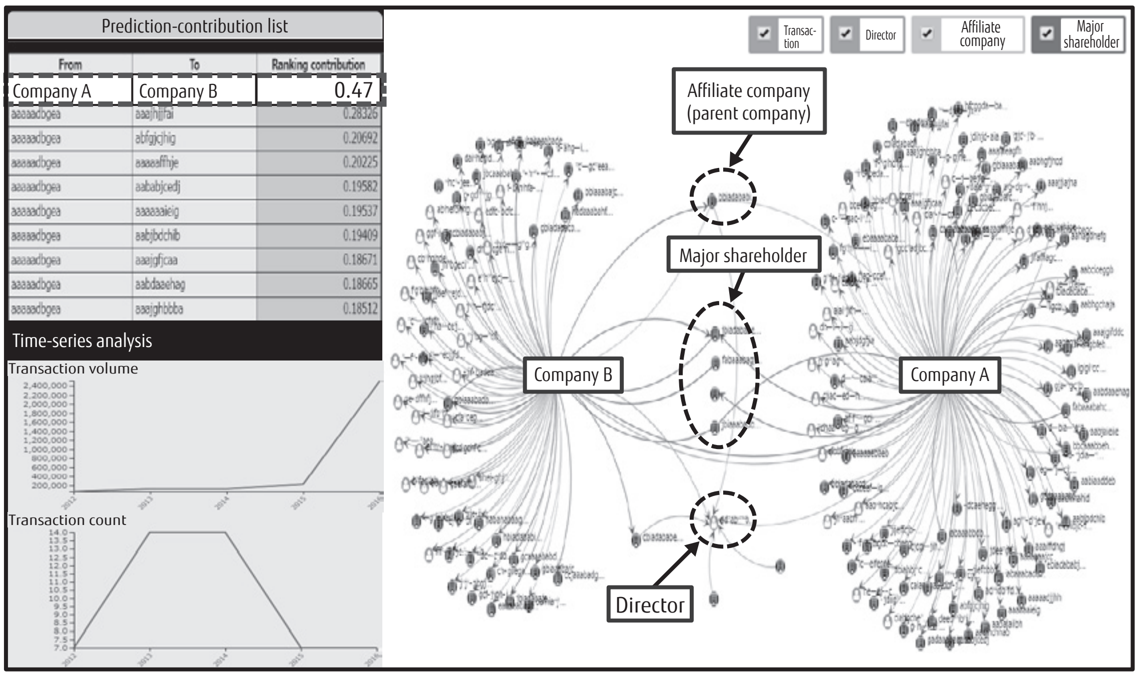


Figure 3
Analysis screen of financial knowledge graph.

risk of company A using this tool. To begin with, it can be seen from the list at the upper left of Figure 3 that exchange transactions from company A to company B contribute the most to the judgment result from among the inference factors output by Deep Tensor. Next, on analyzing the transaction data over time from company A to company B as shown by the time-series analysis in the lower left of the figure, it can be seen that a remarkable increase in transaction volume recently took place and that the number of transactions are in a stable state. In this way, the validity of the result output from Deep Tensor can be verified.

In addition, using a graphical network that visualizes the relationships between companies provides even deeper background knowledge on the inference factors judged by Deep Tensor to be important. For example, it can be seen that common directors exist between company A and company B, which was deemed to be the most important relationship by Deep Tensor in determining the credit score of company A, and that these companies also share many major shareholders while having the same parent company. Taking all the above into consideration, it can be surmised that corporate activities are stable because of the very strong relationship between company A and company B, the existence of common directors and shareholding relationships, and the positioning of these companies under the umbrella of the same parent company. It can therefore be inferred that company A is a low-risk company as a borrower of the bank.

In this way, with the aim of providing trustworthy and explainable AI to the field of finance, the basis of inference factors judged to be important by Deep Tensor can be explained by using a financial knowledge graph constructed with outside information such as open data. We expect that improvements made to the performance of explainable AI will raise the reliability of AI in the future and lead to its wide application in finance.

5. Application example in chemistry

For most materials, functional expression by a molecule alone is infrequent—functions are usually influenced by higher-order structures in molecular aggregates. In addition, most organic materials are used in the form of composite materials or compositions. For this reason, the design and analysis of materials require an understanding of complex relationships including constituent elements and manufacturing processes. At Fujitsu Laboratories, we are working to enhance the reliability of material design and analysis by explaining the basis for material design guidelines and chemical phenomena obtained by machine learning. This section introduces an example of applying knowledge graphs in the field of chemistry.[7)]

Chemical knowledge graphs developed at Fujitsu Laboratories are constructed on the basis of relationships between various sources of information. These include chemical substance databases such as the Japan Chemical Substance Dictionary[8)] and PubChem[9)] that mainly store information on the names and basic properties of chemical substances, chemistry-related patent gazettes, and publically available product information. In addition, we are improving the reliability of explainable AI by using chemical-compound properties, manufacturing processes, and case studies of chemical reactions as information sources that are extracted from the literature using natural language processing. The aim of chemical knowledge graphs is to exploit the feature of uncovering even deeper relationships across a wide range of data.

In the field of chemistry, functions expected of materials are becoming increasingly diverse and advanced, but they are limited by conventional material development techniques that are dependent on the experience and intuition of individuals. As an alternative, attention is now being focused on discovering the factors influencing functional expression from large amounts of data using AI, a process called materials informatics. However, many research themes are faced with problems in this approach, such as the inability to prepare a sufficient amount of data for analysis by machine learning or low reliability in the material design result derived from simulation or machine learning.

To alleviate this problem, Fujitsu Laboratories is using chemical knowledge graphs to supplement a small amount of data with diverse relationships and similarities that cross field boundaries in regards to chemical compounds, properties, compositions, applications, etc. with the aim of improving explainability in material design and analysis. The reliability of material development can be improved by adding features describing the relationships among manufacturing processes and reaction mechanisms instead of relying solely on correlations between simple molecular structures and properties.

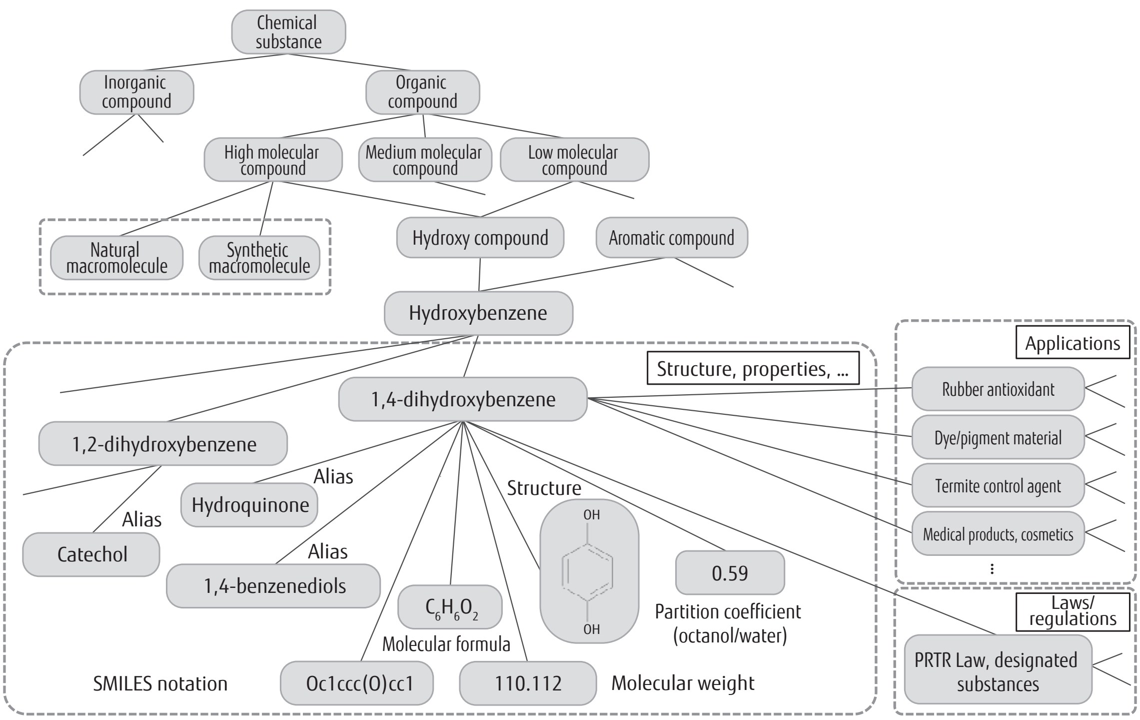
An example of the structure of a chemical knowledge graph is shown in **Figure 4**. To represent the relationships among chemical substances, a group of compounds having the same type of functional group or similar properties can be connected as ontology on a hierarchical structure. To give a specific example, the compound 1,4-dihydroxybenzene connects to the structurally higher-level concept of hydroxybenzene, which is the general name for a compound group that replaces benzene with a hydroxy group as a basic skeleton. Furthermore, the plan is to construct complex graphical networks by making associations from various viewpoints such as hydroxybenzene structure, additives, raw materials, applications, properties, and laws and regulations.

In the above way, it becomes possible to achieve not only a graphical structure of organic compounds but also one that includes the features and phenomena of related compounds. This, in turn, makes it possible to infer the features necessary for material design and analysis even from a small amount of information on properties. In addition, given the large number of synonyms in the notation of organic compounds, making identity judgments between databases while taking ontology into account enables chemical substances to be searched for in patents, journal papers, etc. without being missed. Moreover, by learning the relationships among reaction mechanisms from chemical knowledge graphs in accordance with certain rules, Fujitsu Laboratories aims to achieve efficient searching and automatic design of materials supported by accumulated knowledge.

6. Conclusion

This paper provided an overview of knowledge graphs that form the basis of trustworthy and explainable AI and introduced examples of constructing knowledge graphs in the fields of finance and chemistry.

Amid the R&D of explainable AI now being performed on a global scale, Fujitsu Laboratories has



SMILES: An abbreviation for simplified molecular input line entry system, a method for notation that converts the chemical structure of a molecule into an alphanumeric string with little ambiguity in describing that structure.

Figure 4
Example of a partial structure of chemical-compound knowledge graph.

achieved trustworthy and explainable AI that improves the reliability of AI by providing a reason or basis for an AI inference on the basis of the expert knowledge expressed by a knowledge graph. The implementation of trustworthy and explainable AI in society is now moving forward. From here on, Fujitsu Laboratories intends to refine this technology through R&D focused on actual application fields and to provide "trustworthy and explainable AI" having a level of reliability expected by a variety of fields.

References

1) K. Maruhashi: Deep Tensor Eliciting New Insights from Graph Data that Express Relationships between People and Things. Fujitsu Sci. Tech. J., Vol. 53, No.5, pp.26–31 (2017).
https://www.fujitsu.com/global/documents/about/resources/publications/fstj/archives/vol53-5/paper05.pdf
2) S. Okamoto et al.: AI that spells out its thinking process pioneers a new stage in genomic medicine. BIO Clinica, Vol. 34, No. 3, pp. 48–54, Mar. 2019 (in Japanese).
3) M. Fuji et al.: Explainable AI Through Combination of Deep Tensor and Knowledge Graph. Fujitsu Sci. Tech. J., Vol. 55, No. 2, pp. 58–64 (2019).
https://www.fujitsu.com/global/documents/about/resources/publications/fstj/archives/vol55-2/paper14.pdf
4) S. Tani: Life Cycles of AI-powered Scoring Models Making Use of Transaction Data. FUJITSU, Vol. 70, No. 3, pp. 41–46 (2019) (in Japanese).
https://www.fujitsu.com/jp/documents/about/resources/publications/magazine/backnumber/vol70-3/paper07.pdf
5) Japanese National Tax Agency: Corporate Number Publication Site.
https://www.houjin-bangou.nta.go.jp/en/
6) Japanese Financial Services Agency: Electronic Disclosure for Investors NETwork (EDINET) (in Japanese).
http://disclosure.edinet-fsa.go.jp/
7) K. Tanaka et al.: Knowledge Graph Construction for Chemical Data Utilization Service. Japio YEAR BOOK 2018 (in Japanese).
8) J-GLOBAL Chemical Technology General Link Center: Japan Chemical Substance Dictionary Web (in Japanese).
https://jglobal.jst.go.jp/info/nikkaji
9) NCBI PubChem BioAssay.
https://www.ncbi.nlm.nih.gov/pcassay

Masaru Fuji
Fujitsu Laboratories Ltd.
Dr. Fuji is currently engaged in AI-related research such as knowledge graphs and natural language processing.

Katsuhito Nakazawa
Fujitsu Laboratories Ltd.
Dr. Nakazawa is currently engaged in construction and use of financial knowledge graphs and research and development of AI-quality-control technology.

Hiroaki Yoshida
Fujitsu Laboratories Ltd.
Mr. Yoshida is currently engaged in research and development of chemical knowledge graphs and machine learning.