Knowledge Model for Disaster Dataset Navigation

Yun-Young Hwang*  
Department of Intelligent Data Research, Korea Institute of Science and Technology Information, Busan, Korea  
E-mail: yyhwang@kisti.re.kr

Jin-Hee Yuk  
Department of Intelligent Data Research, Korea Institute of Science and Technology Information, Busan, Korea  
E-mail: jhyuk@kisti.re.kr

Sumi Shin  
Department of Intelligent Data Research, Korea Institute of Science and Technology Information, Busan, Korea  
E-mail: sumi@kisti.re.kr

ABSTRACT

In a situation where there are multiple diverse datasets, it is essential to have an efficient method to provide users with the datasets they require. To address this suggestion, necessary datasets should be selected on the basis of the relationships between the datasets. In particular, in order to discover the necessary datasets for disaster resolution, we need to consider the disaster resolution stage. In this paper, in order to provide the necessary datasets for each stage of disaster resolution, we constructed a disaster type and disaster management process ontology and designed a method to determine the necessary datasets for each disaster type and disaster management process step. In addition, we introduce a method to determine relationships between datasets necessary for disaster response. We propose a method for discovering datasets based on minimal relationships such as "isA," "sameAs," and "subclassOf." To discover suitable datasets, we designed a knowledge exploration model and collected 651 disaster-related datasets for improving our method. These datasets were categorized by disaster type from the perspective of disaster management. Categorizing actual datasets into disaster types and disaster management types allows a single dataset to be classified as multiple types in both categories. We built a knowledge exploration model on the basis of disaster examples to ensure the configuration of our model.

Keywords: dataset relationships, process-based relationships between datasets, dataset visualization, dataset navigation, knowledge model for dataset navigation, knowledge exploration modeling
1. INTRODUCTION

The amount of information and data useful for problem solving is increasing dramatically; however, integrated management systems are insufficient to address this growth because of the lack of interworking systems for data standardization and the increasing introduction of unstructured data. Additionally, multidimensional analyses are difficult because of quality deterioration, including inconsistent data item management by individual institutions and data errors in usage information. With the rapid development of information and communications technology, the public’s proximity to data and interest in the value of data are also increasing. We are rapidly developing into a data-driven society with the use of data-based communication and artificial intelligence technology. Because of the development of data technology and market expansion, various entities, such as disaster-related institutions; medical, research, and government agencies; and local governments, are now using data to solve social problems. Consequently, the demand for data experts is increasing rapidly and the supply of data experts, such as data scientists and analysis experts, is insufficient. As the importance of public data increases globally, the development of information technology is progressing via the establishment of a system of collecting, operating, managing, and sharing disaster safety data to support disaster situation management. As such, research on the use of disaster safety information using artificial intelligence and big data is being actively conducted.

In an environment where there are multiple diverse datasets, there should be an efficient method to provide the datasets needed by users. To address this gap, necessary datasets should be selected on the basis of the relationships between the available datasets.

Presently, a large number of environmental problems, including water pollution, air pollution, global warming, and biodiversity reduction are being encountered worldwide. These environmental problems are regarded as environmental crises because they result in serious damage to human society and nature. The necessary data should quickly and easily be identified to address such disasters. Thus, in this paper, we propose a visualization method on the basis of the relationships between disaster-related datasets. Our method makes it easy to find the datasets necessary for disaster resolution according to disaster type. The relationships between the datasets are derived according to the metadata, dataset field, and dataset value levels. To derive the relationships between the datasets according to the dataset field name and dataset value levels, we built a common knowledge dataset.

The metadata content of the dataset does not include information necessary for the disaster resolution step. For this reason, it is difficult to find a relationship between datasets using dataset metadata. So, we found the relationship based on the field name and data value of the dataset. We derived the relationship between the field names through the standardization of the field names, and the relationship between the data values based on the common knowledge dataset.

This paper is organized as follows. In Section 2, we analyze the related work about knowledge map construction and visualization. In Section 3.1, we describe the process about collection and analysis of datasets. We describe the design of the disaster management ontology in Section 3.2. The disaster management process ontology comprise a process entity and a viewpoint entity, and the viewpoint entity is expressed using the 5W1H methodology. Section 3.3 describes the disaster fields used in the viewpoint entity “what::Type,” and Section 3.3 describes the disaster management ontology. Additionally, this paper introduces a method for identifying relationships between datasets on the basis of common spatial datasets, common temporal datasets, and common measurement datasets. In Section 4, we provide a concise and precise description of the experimental results and their interpretation, as well as the experimental conclusions that can be drawn under the given heading. In Section 5, we describe the ontology evaluation results and how they can be interpreted from the perspective of previous studies and of the working hypotheses.

2. RELATED WORK

To summarize, the research work in this paper helps to understand:

- The need for a process ontology applicable to disaster resolution
- The need for a process-based ontology to understand relationships between disaster datasets
- The need for process views for disaster resolution

2.1. Knowledge Map Construction

Traditional knowledge exploration development methods understand the knowledge of concepts that have the same characteristics as regular documents (Pardos & Nam, 2018). Typically, electronic document management
systems use a theme-based hierarchical classification system to store information. However, in this case, knowledge exploration based on knowledge cross-reference relationships is not implemented. Hence, there is no guarantee that the information that applies to a variety of situations will be utilized. In other words, the hierarchical classification and preservation system does not reflect the characteristics of the information that can be interpreted and applied in various ways depending on the situation. Information can have different meanings depending on the situation. Accordingly, the same information can be applied to various problem situations. To classify and store information to reflect these characteristics, it is more appropriate to use a form of a classification system that can express network types or higher-order relationships rather than a hierarchical classification (Lin & Yu, 2009). Additionally, on the basis of these relationships, the relationships between multidimensional information and mutual exploration functions must be integrated. A knowledge exploration model should be constructed in the form of a network. Thus, a navigation model should be defined and designed on the basis of detailed examinations by experts in information systems, knowledge engineering, and knowledge fields.

Ontology provides a framework that enables effective and efficient knowledge sharing through the formal modeling of a specific concept domain (Gruber, 1995). Using ontology, terms and constraints that describe the structure and behavior of an organization can be defined (Fox & Gruninger, 1998). An organization’s activities, processes, information, and resources can also be modeled (Noy et al., 2000). Ontological techniques can provide a very effective method to describe the relationships and characteristics of related information. Hence, a knowledge exploration model that expresses cross-referenced knowledge relationships using ontology and implements them can expect a synergetic effect that reinforces the essential strengths of the two techniques. In other words, information has no independent meaning, and the meaning of information can become clear depending on what purpose and in which process we use it. Thus, the ontological class and the “knowledge-purpose-process” relationship can easily be defined via a definition procedure. Additionally, when inferring the relationships between concepts defined in an ontology, new information and relationships that are not already included in the knowledge exploration model can be identified. Consequently, the knowledge exploration model can be automatically extended. Accordingly, multiple studies can be found highlighting the need to apply ontological techniques when building a knowledge management system (Savvas & Basiliades, 2009). However, most existing studies are presently only at the level of formally expressing information and related concepts and do not consider cross-reference navigation on the basis of network relationships between different pieces of information.

The accuracy of knowledge retrieval and extraction by users is determined by how information is categorized and managed according to the storage system. Storing information on the basis of simple titles and keywords can improve storage efficiency. Nevertheless, the applicability of a given piece of information is difficult to ensure. Information must be classified and stored according to the process or purpose in which it is stored and relation to the process to which the extracted information is applied, ensuring the practical usefulness of the search (Kim et al., 2003).

Process-based knowledge exploration shows the current state of knowledge utilized by each process, depending on the intended use of the process (Abecker et al., 2001; Lee et al., 2006). Some information is associated with multiple processes or purposes, such information forms a network type of relationship with other information. As can be seen here, to express the current state of knowledge, it is desirable to regard a piece of information as a relational object linked to another piece of information rather than as an independent entity. Accordingly, the concept of ontology can be used to express the meaning of the relationships between pieces of information (Xiong et al., 2008). In other words, to express all the meanings of a piece of information in the real world, it is essential to define its relationship with the ontology-based knowledge concept. Based on this, knowledge exploration can spontaneously infer and expand these relationships with other knowledge concepts.

Malone et al. (2003) introduced the method of constructing a process ontology using the MIT Process Handbook. An ontology created by following the guidelines of the MIT Process Handbook can express a large number of processes and can express a special hierarchy of processes and interrelationships of properties. All major parts of the MIT Process Handbook, including processes, bundles, targets, exceptions, resources, dependencies, and tradeoffs, are represented as OWL classes (Mu et al., 2018).

In most cases, relevant studies that have proposed process ontology-based knowledge exploration models have been applied to product handling processes (AbdEllatif et al., 2018; Boubaker et al., 2012; Köpke & Su, 2015; Mu et
al., 2016; Rao et al., 2012; Zhang et al., 2017). These studies have focused on reducing the error rate of the product handling process (Boubaker et al., 2012; Köpke & Su, 2015; Rao et al., 2012; Zhang et al., 2017) or guiding process guidelines and procedures (AbdEllatif et al., 2018). However, Mu et al. (2016) had been made to simplify knowledge retrieval rules using a process-based ontology. To provide a cooperative environment for various stakeholders, they propose a methodology for collaboration between various process-based ontologies. They propose transformation rules to transfer the collected process-based ontology to the mediator concept (transferred knowledge).

2.2. Knowledge Map Visualization

Data visualizations are the best way to show users the relationships of datasets. We investigated studies on topics such as temporal data visualization, geographic data visualization, linked data visualization, and ontology visualization.

For visualization specifications, the related papers introduced a survey about classification, data source, presentation medium, etc., of visualization languages (Mei et al., 2018). We surveyed visualization languages from the stack perspective and emphasize how these languages are used from a practical perspective. There have also been some surveys on exploratory data analysis tools (Diamond & Mattia, 2017; Ghosh et al., 2018), which are complementary to our interactive data visualization.

For efficient approaches for data visualization, the approaches consider how to integrate databases, data visualization, and data analysis so that a user can easily work in one system, but without a discussion for efficiency (Keim et al., 1995). Idreos et al. (2015) surveyed the techniques which aim to improve efficiency in the data exploration cycles, but we focus on techniques about how to construct visualizations efficiently. Bikakis (2018) gives an overview of current systems and techniques for big data visualization, but with a less detailed discussion.

For data visualization recommendation, although there have been many works about recommendation systems (Adomavicius & Tuzhilin, 2005; Bobadilla et al., 2013; Burke, 2002; Sharma & Gera, 2013) and works about recommendation for different tasks, e.g., quality of service (QOS)-aware web services (Christi & Premkumar, 2014), social software (Guy et al., 2009), and E-commerce (Wei et al., 2007), there is no survey about data visualization recommendation.

Of the above visualization methods, we used a compromise between temporal data visualization (Wohlfart et al., 2008) and linked data visualization (Coimbra et al., 2019; Dadzie & Pietriga, 2017; Marie & Gandon, 2014).

![Fig. 1. Process to build a knowledge model for knowledge navigation. API, application programming interface; DB, database; LOD, linked of data.](https://doi.org/10.1633/JISTaP.2021.9.4.3)
The linked visualization method is primarily used to express relationships between datasets. Such methods show data relationships in radial or hierarchical forms. Because we must visualize the relationships between the datasets, those between the dataset and the business process, and those between the dataset and the business type, there is a limit to our ability to use the graph visualization method.

The temporal data visualization method can illustrate a data relationship diagram that changes over time. Accordingly, we used the temporal data visualization method because the required dataset is different depending on the time of the disaster response.

3. MATERIALS AND METHOD

3.1. Collection and Analysis Dataset Process

In this section, we describe the entire process including dataset collection, dataset analysis, and extract association between dataset. As shown in Fig. 1, to build a knowledge navigation model, we (i) collected and analyzed datasets, (ii) coordinated the datasets based on the field names that made up the datasets, and (iii) built an ontology based on the standard field names.

We collected all the datasets from the Korea Public Data Portal (by the Ministry of the Interior and Safety, n.d.). Among the collected datasets, we found the linkage between the datasets by targeting only the structured datasets. A structured dataset refers to a dataset consisting of pairs of field name and data value. To increase the connection between the datasets, we analyzed the values of the raw data and extracted common field names in the dataset. We analyzed the field names of raw data to define standard field names. Then the associations between the datasets were defined based on the defined standard field names.

We only considered “isA,” “sameAs,” and “subclassOf” relationships in this paper. We stored all datasets and their relationships in a relational database. This method has a limitation in which we can express only simple relationships. To increase the connections between the datasets, we analyzed the values of the raw data and extracted common field names, as shown in the Fig. 1. This paper describes in-depth how to connect the collected data sets based on process ontology.

3.2. Ontology for Disaster Management Processes

One dataset must be matched to the disaster management process. To map the dataset and the disaster management process, we first matched the dataset used as the input for the tools utilized by the domain experts in the disaster management process. To verify this match, a domain expert was consulted to check the results.

We designed the disaster management ontology in reference to the MIT Process Handbook. The disaster management process ontology comprises a process entity and a viewpoint entity. Fig. 2 shows our knowledge navigation model. Our knowledge navigation model is organized by concept, process, and viewpoint. In this model, there is a hierarchical relationship between the concepts and there is a context between the concepts depending on the process-

![Fig. 2. Process-centric knowledge navigation model.](http://www.jistap.org)
es. In addition, processes and concepts can be expressed via various relationships depending on the viewpoint.

The perspective used in our model incorporates $V_{sw1H}$ from a disaster perspective. In particular, $V_{sw1H}$ retains the six elements of disaster type, cause of disaster occurrence, management method, time of occurrence, location of occurrence, and organization. In addition, the mapping is coupled such that why::Cause, what::Type, where::Location, when::DateTime, who::Organization, and how::Method.

The rules for configuring this ontology are as follows:

- **P: Process entity for disaster management**
- **$V_{sw1H}$: View entity retaining the following six elements: the disaster type, cause of the disaster occurrence, management method, time of occurrence, location of occurrence, and organization**

  a) **why::Cause** — This viewpoint represents the cause of the disaster occurrence.

  b) **what::Type** — This defines the type of disaster. The disaster types are described in detail in Section 3.2.

  c) **where::Location** — This viewpoint classifies the location information related to where a disaster occurred and disaster-related location information, such as evacuation shelters. Depending on how the disaster is managed, information concerning the locations of the occurrence and/or shelters is required.

  d) **when::DateTime** — This expresses the time of the disaster occurrence and the time of the disaster management. That is, this viewpoint is used to classify related datasets according to time.

  e) **who::Organization** — This relates to information concerning disaster-related data producers and disaster management organizations.

  f) **how::Method** — This represents the Korean disaster

![Disaster Categories in the Republic of Korea](https://doi.org/10.1633/JISTaP.2021.9.4.3)
management classification. There are four disaster management standards: mitigation, preparedness, response, and recovery. We describe the disaster response measures in detail in Section 3.3.

The relationships between the process, viewpoint, and information can be expressed via the following conceptual formula:

\[ P_i = f(P_j \| V_j \| D_j) \]

Here, \( P_i \) represents the process \( i \), where \( i = 1, 2, 3, \ldots, n \); \( V_j = g(P_k) \), where \( V_j \) represents the view \( j \) composed of the process \( k \), where \( j = 1, 2, 3, \ldots, n \) and \( k = 1, 2, 3, \ldots, n \) and \( D_j \) represents the dataset \( j \) used in the process \( i \).

3.3. Ontology for the Disaster Categories
This section describes the disaster field used in the viewpoint entity ‘what::Type’ described in Section 3.1.

In Korea, disasters are managed in three categories, as shown in Fig. 3, according to the Misfortune and Safety Supervision Basic Law. Natural disasters refer to disasters caused by natural phenomena, such as typhoons, floods, heavy rain, storms, tsunami/storm surges, snowfall, drought, earthquakes, yellow dust, and red tides. Man-made disasters correspond to damages greater than the scale determined by the associated presidential decree and have subclasses of fire, explosion, collapse, traffic accidents, chemical-biological-radiological accidents, and environmental pollution accidents. Social disasters refer to damage caused by a paralysis of the national infrastructure, e.g., energy, communications, transportation, finance, medical care, or water supply, or the spread of infectious diseases.

3.4. Ontology for Disaster Management Categories
Disaster management is defined as activities performed to minimize the damage due to a disaster. In the case of the Republic of Korea, disaster management is divided into mitigation, preparedness, response, and recovery stages in the Misfortune and Safety Supervision Basic Law (Fig. 4).

According to the United Nations International Strategic Organization for Natural Disaster Reduction, the

Fig. 4. Process to build a knowledge model for knowledge navigation.
mitigation phase refers to the mitigation of the side effects of a disaster in relation to the risks. The preparedness stage includes disaster training to prepare for disasters and for damage caused by disasters and activities designed to increase public awareness. The response stage provides necessary resources to the victims after an emergency or disaster occurs. The recovery stage comprises all activities designed to reduce damages due to a disaster. This step, in particular, focuses on efforts to restore communities affected by the disaster to their normal state.

We built an ontology for the four stages described above. This information is used as a classification system corresponding to the “how::Method” of the disaster management process ontology described in Section 3.1.

3.5. Defining Relationships between Ontologies and Datasets

As shown in Fig. 5, datasets must match the disaster category information and the disaster process information. We used the metadata of the datasets to map the datasets and disaster category information. We first matched the datasets used in the flood, fine dust, and earthquake prediction tools to map the datasets and disaster field processes and then verified the result with the experts. The relationships between the datasets can be found by connecting common knowledge datasets.

In other words, our methodology allows the processes required for disaster response to be determined according to the type of disaster and maps the dataset required for these processes to provide the necessary datasets according to the disaster response situation.

4. EXPERIMENTAL RESULTS

4.1. Building Dataset for Experiment

We stored all collected datasets in MySQL database. We stored dataset metadata, dataset field names, and relationships between the dataset in the database (Fig. 6). Since we express the relationship between datasets using only the “isA,” “sameAs,” and “subclassOf” relationships, the relationships between datasets can be stored and managed in a database table. Since these are very basic relationships, we need to derive high-level relationships using inference on linked data (LOD) data built with resource description framework (RDF) triples.

We assigned a primary key to each dataset, and each dataset was linked to a “Field names of dataset” table and a “Relationship between dataset based on ontologies” table. The “Relationship between datasets based on field names exact matching” table has information derived from the relationality through standardization of field names, and the “Relationship between datasets based on ontologies” table has the relation between the common knowledge dataset and the actual values of the dataset.

We defined the relationship between the tables represented in the figure as follows.

- subclassOf: relationship between metadata and field names of dataset
- isA: relationship between relationship tables, which are “relationship between datasets value based on common knowledge dataset,” “relationship between datasets based on field names exact matching,” and

![Fig. 5. Relationships between our ontologies and datasets.](image-url)
Yun-Young Hwang, et al., Knowledge Model for Disaster Dataset Navigation

“relationship between datasets based on ontologies,” and other tables
• sameAs: relationship between “standard field names” and “similar field names”

Table 1 shows the result of converting the data set instance stored in the database, the disaster field category, the disaster management process, and the relationship between them into RDF triple.

To convert information stored in the database into RDF triples, we used D2RQ (Accessing Relational Databases as Virtual RDF Graphs) open source platform. In addition, we have complied with the W3C Recommendation standard, A Direct Mapping of Relational Data to RDF guidelines. We developed our ontology verification function using SHACL4P within Protégé (Fig. 7). First, with the help of Jena API, we extract all the information about ontological constructs, e.g. classes, sub classes, object properties, data type properties, their domain and range, data types, and restrictions, according to the mapping algorithm. After mapping classes and subclasses into relational database format, we transformed object type properties according to defined mapping rules. According to mapping rules, we then transform single valued and functional object type property into the foreign key in the table that relates to domain of object type property and this key reference to primary key in the table that relates to the range class of object type property.

4.2. Developing a Common Knowledge Dataset

To increase the connections between the datasets, we analyzed the values of the raw data and extracted common field names. We calculated the connections between the datasets on the basis of the relationships between the dataset field names and the similarities between the raw data values. We collected approximately 30,000 datasets and extracted all the field names in each dataset. We extracted approximately 1.8 million field names and found several similar field name groups. As a result of analyzing the field names, the space-related field names accounted for 35% of the total. It was followed by time-related field names (17%), measurement information-related field names (15%), and

<table>
<thead>
<tr>
<th>Table 1. The number of RDF triples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of records in dataset</td>
</tr>
<tr>
<td>110,555,727</td>
</tr>
</tbody>
</table>

RDF, resource description framework.
data set management-related field names (14%). We constructed several common knowledge datasets to derive the relationships between the different field names belonging to each field name group. In this paper, we introduce a method to discover relationships between datasets on the basis of our common spatial, common temporal, and common measurement datasets.

As shown in Fig. 8, we derived relationships between the datasets on the basis of the common spatial datasets. We defined the “road name address” as a representative value of the common spatial datasets; this value was designed to enable mapping of the coordinate value, latitude/longitude, lot address, mailbox address, and administrative area around a road name address. The common spatial datasets allowed datasets composed of different field names to be linked around road name addresses.
4.3. Identifying Method-Related Datasets

To determine the datasets related to the query described in Section 3.5, we expressed the following relationship for each dataset as a matrix:

\[ RD_i^* = (\text{DisasterType}, \text{Preparedness}, \text{Mitigation}, \text{Response}, \text{Recovery}) \]

Here, \( RD \), represents the relation dataset \( i \), where \( i = 1, 2, 3, \ldots, n \), and \( \text{DisasterType} \) is the type of disaster, e.g., fine dust, flooding, or earthquake.

When \( RD \) can be used in the disaster management process, its value is 1; otherwise, it is 0. For example, \( RD_1 = (\text{Flooding}, 1, 0, 1, 0) \) means that \( RD_1 \) is a flooding-related dataset and can be used for “preparedness” and “response” during the disaster management process. Because a single dataset can be used for multiple disaster types, multiple \( RD \) datasets can exist.

Fig. 9 shows an example of the relationships between an instance built according to the above described ontology and an actual dataset. The gray region in the figure panel indicates the ontology instance area, and the other regions represent the actual data values.

To confirm that the configuration of the knowledge navigation model is correct, we defined the following query and a statement of quality suitable for our knowledge navigation model. We expect the following result for the above query in the figure.

\[ \text{Datasets} \rightarrow \text{Flooding mark, Magnitude, PM2.5 (particulate matter with a diameter of less than 2.5 um), Rainfall, Wind Speed, Population, Wind Direction} \]

\[ \text{Query: Search for datasets needed for flood forecasting in Busan, South Korea} \]

\[ \text{Redefined Query: Find (dataset, Cause, Type, Location, Time, Organization, Method)} \]

\[ \text{Result: Flooding Mark, Rainfall} \]

\[ \text{Defined Scenario-Based Query: Model-based V5W1H} \]

\[ \text{why::Cause = \{all\}} \]
\[ \text{what::Type = \{flood\}} \]
\[ \text{where::Location = \{Busan\}} \]
\[ \text{when::DateTime = \{all\}} \]
\[ \text{who::Organization = \{all\}} \]
\[ \text{how::Method = \{Prevention\}} \]

4.4. Visualization of the Relationships between Datasets

We developed a dataset navigation prototype on the basis of the dataset navigation model.

The red, yellow, and blue bands shown in Fig. 10 indicate the type of disaster, and the black triangles, squares,
and circles indicate the disaster management processes. The relationships between these datasets are represented by the interior lines. If the line connecting two datasets is thick, it means that the relationship between these datasets is strong. Stronger connections between datasets also indicate that the data sources of the datasets are different.

5. ONTOLOGY EVALUATION

To evaluate our methodology, we conducted a usability evaluation on the knowledge navigation prototype with a group of domain experts in the fields of air pollution, floods, and earthquakes.

First, we conducted a survey to see if our ontology can help users discover the dataset they need. We asked the domain experts two questions. The first question is whether the dataset navigation described in Section 4.4 allowed experts to discover the data set they wanted. This was to confirm that the disaster-related classification and disaster-related process ontology were well established, and the relationship between the two was clearly expressed. Fig. 11A shows the satisfaction score for our methodology. As can be seen in Fig. 11B, we obtained an average of 5.8/7 points for the appropriateness of the knowledge modeling method. The earthquake expert gave the lowest score, whereas the fine dust expert gave the highest score. This score was derived because our dataset was built with a bias towards the fine dust dataset.

In order to compare the dataset search before and after applying the process-based knowledge search model, we asked two groups to search with the same query word before and after applying the model. To increase confidence in the results, the experts were divided into two groups: Groups A and B. These groups consisted of experts in the fields of air pollution, floods, and earthquakes. We compared the navigation before and after the application of the process-based knowledge navigation model. Group A used a knowledge navigation system to which our methodology was not applied, whereas Group B utilized a knowledge navigation system to which our methodology was applied. The experts used the two systems to answer the same questions. We then measured the time taken by the disaster workers to find the datasets they needed. Knowledge exploration reduced the data retrieval time on average by 73%, that is, the average search time with our methodology was 27% higher than without our methodology. To measure the efficiency of the model, we calculated the proportion of the datasets that were meaningful to the total identified datasets. Table 2 shows that using our model resulted in a lower efficiency. We found that the reason for these results is that our knowledge search model consists of only simple relationships such as isA, sameAs, and subclassOf. We plan to improve our model to resolve this issue.

![Visualization of relationships between datasets.](image.png)

**Fig. 10.** Visualization of relationships between datasets.

![Satisfaction distribution score](image.png)

**Fig. 11.** Result of the survey. (A) Satisfaction score for our methodology and (B) average satisfaction scores for the various domain experts.
6. CONCLUSION

A knowledge exploration model comprises diagrams that can be understood in the form of a network; this not only allows the current state of knowledge to be checked at a glance but also comprises cross-reference relationships with the network knowledge, implemented via exploration and related information. It is a good idea to improve the accuracy of exploration models and to make it easier to search through and work with various datasets. In this study, a procedure to expand the original meaning and function of a knowledge search model was developed by fusing the function of the ontology with the existing knowledge search.

The relationships between pieces of information included in the knowledge exploration model covered in this study were defined according to the datasets required by the tools used for disaster resolution. However, managing numerous types of information requires an automated process to understand the content of the information and relationships between other concepts and to construct a knowledge exploration model. Thus, in the future, we will perform automated knowledge acquisition to interpret the content and relationships of acquired information and build an automated knowledge exploration model. Automated knowledge acquisition can be achieved via the monitoring and interpretation of source conversations, documents, and actions that contain information. Understanding the content and relationships of collected information is possible via the development of techniques to extract titles and key controls from documented content.

Using our proposed method, a flexible relationship between datasets can be built according to the purpose and cause of problem solving, allowing users to discover the datasets necessary to solve a specific problem. In addition, the user can grasp the importance of the data set by determining the time, space, and type of problem.

In the future, we will apply this methodology to a larger number of experimental datasets. Additionally, we will perform automatic knowledge acquisition to understand the content and relationships between the acquired information and to build an automatic knowledge exploration model. Additionally, we plan to continue research on deriving relationships through reasoning.

ACKNOWLEDGMENTS

This research was conducted with the support of the Open Data Solutions (DDS) Convergence Research Program funded by the National Research Council of Science and Technology “Development of solutions for region issues based on public data using AI technology—Focused on actual proof research for realizing a safe and reliable society.”

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES


Köpke, J., & Su, J. (2015, August 31-September 3). Towards ontology guided translation of activity-centric processes to GtM. In M. Reichert, & H. Reijers (Eds.), *Proceedings of the BPM 2015, 13th International Workshops* (pp. 364-375). Springer.


