# Computer-assisted Structuring of Emergency Management Information: A Project Note

**Yotaro Watanabe    Kentaro Inui**
Tohoku University

{yotaro-w,inui}@ecei.tohoku.ac.jp

**Shingo Suzuki**
Kyoto University

shingo@drs.dpri.kyoto-u.ac.jp

**Hiroko Koumoto**
Fuji Tokoha University

koumoto@fuji-tokoha-u.ac.jp

**Mitsuhiro Higashida**
ESIP

m-higashida@kansai-kumikomi.net

**Yuji Maeda**
NTT Secure Platform Laboratories

maeda.y@lab.ntt.co.jp

**Katsumi Iwatsuki**
Tohoku University

iwatuki@riec.tohoku.ac.jp

## Abstract

In order to achieve high-level resilience against disasters, effective utilization of previous emergency management information is necessary. The goal of this project is to establish effective utilization of emergency management information and emergency response logs that are accumulated as a fundamental dataset to learn lessons for emergencies in the future. More precisely, we develop a framework that simplifies structuring emergency management information and creating databases through various media or formats by exploiting technologies such as natural language processing to fix the bottlenecks for inputting information in emergency response sites, to share disaster state, and to contribute towards achieving more effective use of human resources. The academic aim of this project is to establish the task of creating a database of emergency management information as a subfield of natural language processing applications.

## 1 Introduction

In order to confront natural disasters which could become a national crisis, high-level resilience against disasters is required. To achieve this, it is necessary to assume emergency situations, prompt actions for emergencies, and conduct quick and correct restorations and recoveries. This requires effective use of emergency management information.

In response to the 2011 Tohoku earthquake/tsunami in Japan, a strong desire for the development of new methods to improve emergency response came about. The basic framework for emergency management by local governments is the following: (i) As soon as a disaster takes place, the local government organizes a headquarter for disaster control, which consists of the mayor (as the director-general) and directors from executive branches (Water and Sewer Division, Civil Engineering Division, etc.) as well as representatives from the police, the fire station, etc. (ii) Each executive branch collects disaster information from a large variety of sources and responds to requests from disaster sites. (iii) The collected information is conveyed to the headquarter and also shared with other branches so that the headquarter and branches can effectively cooperate with each other. (iv) The headquarter keeps the situation in perspective and makes local government-wide decisions. (v) A summary of the accumulated disaster information and responses is communicated to neighboring local governments and the higher administrative division (i.e. prefecture or state).

As the reader may imagine from the above, the most important key for this whole system to work effectively is communication; i.e. the key issue is how efficiently and precisely information about the progress situation and the responses against it can be shared among the executive branches and the headquarter together with the outside of the disaster site. Unfortunately, however, the 3.11 earthquake revealed that in most local governments in the disaster-hit areas, the current means for emergency management communication was crucially inefficient, which sometimes caused miscommunications and prohibited the disaster control headquarters from making optimal or appropriate decisions. While the inefficiency of communication was partly due to severe damages of communication networks, critical problems arose even under the situation where communication networks were available. In most of the local governments in the disaster-hit areas, each piece of information from outside (through phone calls, radio communications, etc.) was recorded only
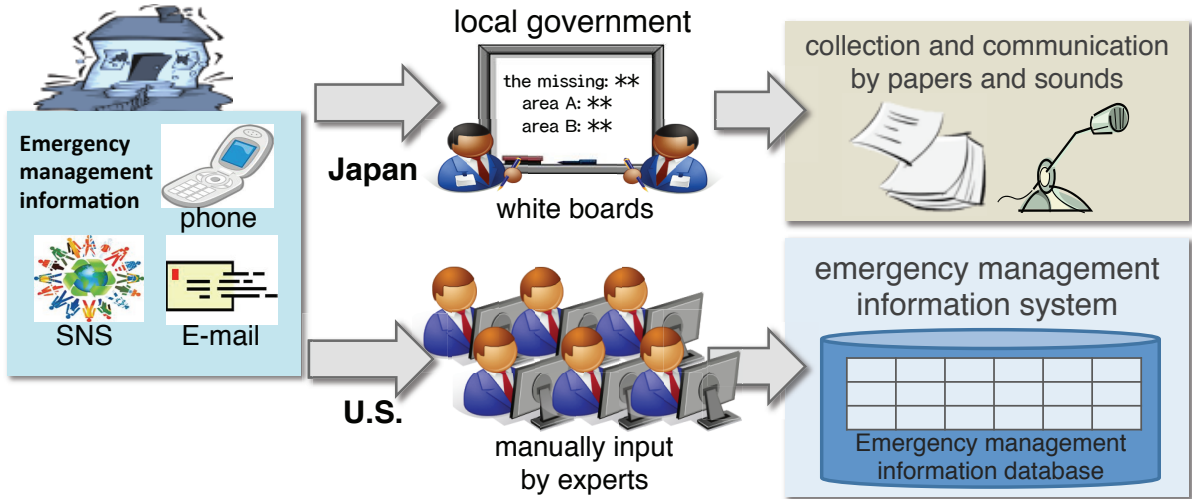
Figure 1: Emergency management systems in Japan and United States.

by hand writing and thus distributed to divisions in charge only by oral communication or through white boards as shown in Figure 1. This makes it difficult to efficiently share emergency management information among divisions and local governments, also hinders promptly update status of disaster response.

In contrast, in the United States, WebEOC (EOC: Emergency Operation Center), a web browser-based information system conformable to a standardization of incident management by NIMS (National Incident Management System), has been introduced to more than 80% of the state governments, and has achieved effective emergency responses. To operate the systems, professional employees are hired at each crisis management office to manually input emergency management information as structured information using a computer. Standardization of information management in crisis situations has internationally progressed and has been recommended for ISO 22320. On the other hand, such emergency management information systems have not been introduced to actual situations in Japan.

It is not enough that simply converting emergency management information from a large variety of sources to unstructured texts. Realizing quick and correct restoration and utilization of them for the future disasters needs making them as structured information to be searchable by storing, classifying and organizing them. The structurization consists of selecting one of database schemas and search its items from texts, and prohibitively

high cost will be required if the operation is performed by human. In order to operate such system with limited resources, reducing the cost of structurization is essential.

Given this background, we have launched a government-founded three-year project to develop a system that assists with creating databases of emergency management information to digitize, accumulate and utilize them. We address the following three issues in this project.

**Designing standard DB schemas for emergency management information** By analyzing actual emergency response operations, we design standard DB schemas for emergency management information that can be commonly applied for diverse local governments.

**Developing a system that assist with creating emergency management information databases** By further advancement of natural language processing technologies, we develop a system that efficiently stores unstructured emergency management information to databases. This system at first extracts important information from unstructured texts, selects one of database schemas, and finally fills each of the elements with a corresponding expression in a text, as shown in Figure 2.

**Environmental improvement for the usage of emergency management information systems** We develop a training package that includes diverse and real scenario data and analyzes issues of developing information and communication technologies for emergency response through experi-
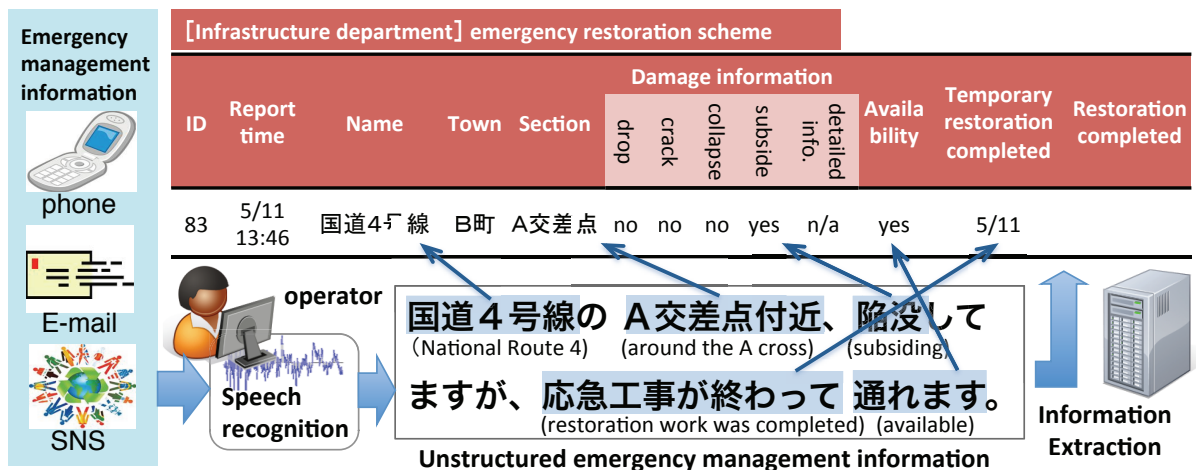
| [Infrastructure department] emergency restoration scheme | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ID | Report time | Name | Town | Section | Damage information | | | | | Availability | Temporary restoration completed | Restoration completed |
| | | | | | drop | crack | collapse | subside | detailed info. | | | |
| 83 | 5/11 13:46 | 国道4号線 | B町 | A交差点 | no | no | no | yes | n/a | yes | 5/11 | |

Figure 2: Creation of emergency management information databases.

ments of technologies of adding emergency management information to databases.

In addition, we develop a system that unifies these research outcomes and evaluates performance of the developed system by conducting experiments cooperated with the emergency management information system. Through discussions of technical and systematic issues we will have, we find knowledge for deploying the system into the field.

## 2 Research Issues and Plans in Our Project

In this project, we develop an emergency management information database creation support system using speech recognition and natural language processing technologies as shown in Figure 3. To do this, we (1) design emergency management information schemas, (2) develop an emergency management information database creation support system whose inputs consist of various forms of information such as speeches, faxes with handwritten characters, and so on, and (3) improve the environment for the usage of the emergency management information system. In addition, we incorporate the developed system into WebEOC, a standard emergency management information system, and then conduct demonstration experiments with the developed system by cooperating with local governments to analyze technical and systematic issues in deploying the system into the field. The detailed explanations of the tasks we address are explained as follows.

### 2.1 Task 1: Designing Emergency Management Information Schemas

In this task, we at first analyze emergency response operations thus far in local governments cooperating in this project. We then design and standardize emergency management information schemas and develop a system that assists with creating emergency management information databases. In addition, we develop a framework that automatically creates emergency management information in conjunction with WebEOC.

In our prior work, we developed a set of emergency management information schemas applicable to earthquake emergencies in local governments. In this project, we develop such schemas applicable for not only local governments, but also umbrella organizations of them, such as ordinance-designated cities and administrative divisions of Japan, and for central governments. In addition, we design a set of standardized schemas applicable for other emergency situations such as wind damage, flood damage, eruption and pandemics. In research and development, we not only clean up elements of emergency management information, but also extract and standardize them to be able to apply for several emergencies and local governments by analyzing results of demonstration experiments. More precisely, we first go to interview employees of disaster affected local governments. Next, we organize disaster response instances in chronological order by referring various forms of histories and analyze them to clear issues regarding emergency information processing. Because there are local governments who man-
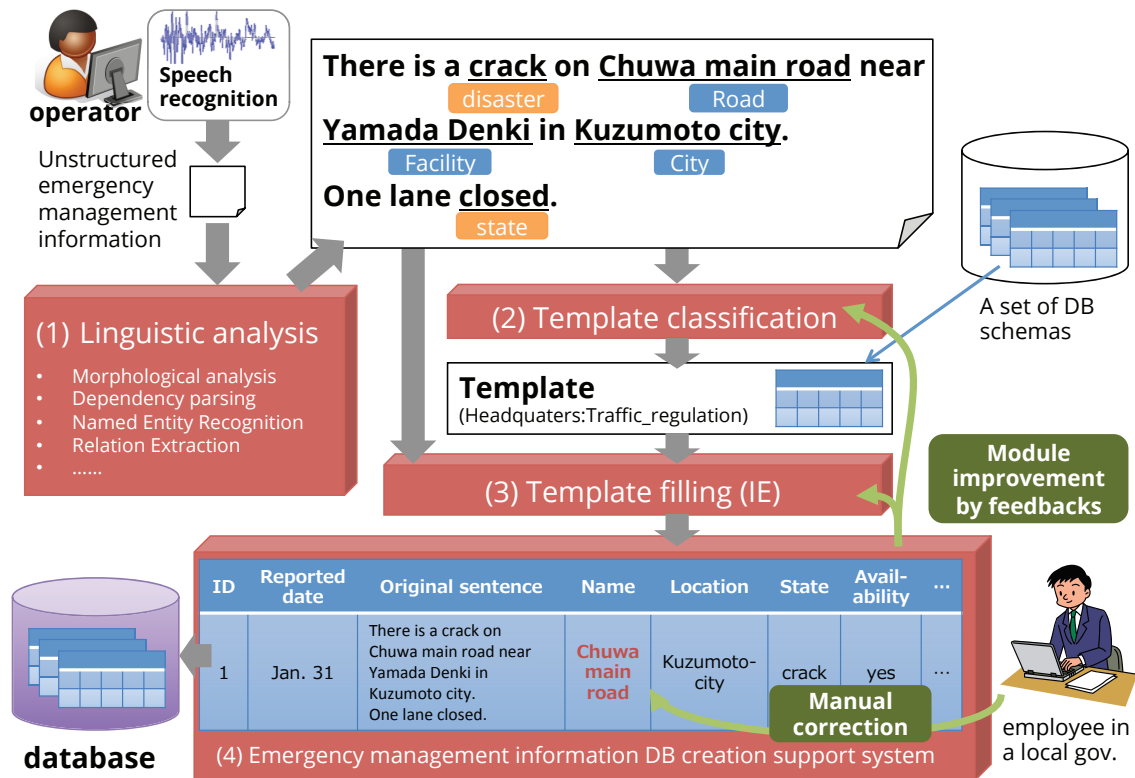
Figure 3: An emergency management information database creation support system.

age pieces of emergency information using paper-based media such as FAX and share them in their own formats as shown in Figure 4, we consider the needs for these local governments by establishing sharable and flexible emergency management information schemas customizable for each local government.

Then, we standardize emergency management information schemas through comparison to outcomes of previous work. Standardization of emergency management information schemas is conducted through several opportunities such as domestic or international conferences, and developed on a cloud-based system. For this, we develop a guideline of the cloud-based system as a emergency management cloud through ASPIC (ASP-SaaS-Cloud Consortium) [1] and encourage broad use of the system.

### 2.1.1 Research Issues and Plan

Toward efficient sharing of emergency responses, United States defined Incident Command System (ICS) Forms, a standard of federal emergency management, in which items required for ICS are defined. The templates of them were provided to

WebEOC. However, it is unclear whether those forms are both necessary and sufficient for operating response activities. Additionally, they may not be applicable to response in Japan. In a Kashihara city case study, Higashida et al. (2012) created operational templates to deal with necessary data. Schemas of the templates indicate necessary data items for operations. They are, however, not based on information in actual response. In order to establish standard emergency management schemas which are applicable for various types of emergency situations and usable in ordinance-designated cities, the administrative divisions of Japan and central governments, we interview local governments to draw out information of past emergencies and analyze them to address the issue.

### 2.2 Task 2: Development of an Emergency Management Information Database Creation Support System

The task of developing an emergency management information database creation support system consists of (a) development of emergency management information structurization technologies and (b) development of a user interface for creating

Figure 4: Different forms of damage reports.

emergency management information databases.

### 2.2.1 Developing Emergency Management Information Structurization Technology

We assume that inputs in the system are digitized unstructured texts transformed from primary emergency management information such as sounds and papers, e-mails, and social networks. We develop a structurization technology that automatically extracts information which corresponds to items in schemas from unstructured texts.

Considering actual use of this technology, we have to handle not only text data but also speeches and images. However, considering the fact that this research has a limited time frame of only three years, we decided to concentrate on creating databases from digitized, unstructured texts. Since speech recognition performance depends on speech environment, acoustic model, etc., it is necessary to consider such factors into improve the accuracy of speech recognition which digitize emergency management information. In this project, we assume that speech input is performed by some particular operators. This enables us to provide invariable environment for speech recognition. Also we consider using a fixed form for reading out emergency management information. On the other hand, existing hand-written character

recognition technologies are currently not reliable for use, so it would be necessary to read handwritten texts out loud and digitize the speech via speech recognition.

Structurizing emergency management information of digitized texts and converting them as databases can be seen as a task of information extraction. However, in contrast to the traditional information extraction tasks, our task is more complicated and challenging because there are diverse types of entries in emergency management information schemas to fill. For example, this task requires handling various types of information: not only named entities but also domain-specific events (subside, fire, etc.), modality information (e.g. available), etc. Since we have already developed several natural language processing technologies, we advance these technologies along with taking measures to adopt them for emergency management information. Also, as shown in Figure 2, various linguistic knowledge and domain knowledge are required to structurize emergency management information. For instance, we have to recognize that *a cross* represents a point in a road, the cross is located in a particular area, and "can pass through" means "the road is available". Thus exhaustive acquisition of such knowledge is critical to the successful development of this technology. We also advance investigating technolo-

gies of large-scale linguistic and domain knowledge acquisition from web pages.

### 2.2.2 Development of a User Interface of the Emergency Management Information Database Creation Support System

For the emergency management information databases we design in this project, we prepare schemas for more than ten divisions of local governments and define dozens of items for each of the schemas. The structurization we described in Section 2.2.1 requires extraction of extremely fine-grained database items from natural language text, and this kind of difficult task setting has not yet been explored in previous work. Thus, instead of using system outputs without change, we need a system which can easily display choices presented by the system for operators to select. To address this, we develop a high-quality framework that efficiently create databases from unstructured emergency management information by effectively using emergency management information structurization technologies. These technologies are improved through interactions with users and machine learning approaches which enable us to dynamically improve the performance of the system by user feedback.

### 2.2.3 Research Issues and Plan

The structurization includes diverse information extraction subtasks including named entity recognition (NER) such as location names (city, road, etc.), facility names (shelter, shop, school, etc.), numerical expression identification and normalizing, relation extraction (RE), location name disambiguation and slot filling.

In Message Understanding Conference (MUC) (Grishman and Sundheim, 1996) and Automatic Content Extraction (ACE) (Doddington et al., 2004) communities, various information extraction tasks including named entity recognition, relation extraction and slot filling have been explored. The task of disambiguation location expressions is called toponymy disambiguation, and has been explored by (Buscaldi and Rosso, 2008b; Buscaldi, 2010; Buscaldi, 2011; Habib and van Keulen, 2013; Bo et al., 2012; Lee et al., 2013) and GeoNLP Project[2]. For the task of disambiguation location names, Buscaldi and Rosso developed Geo-WordNet (2008a) in which entries

---

[2]http://agora.ex.nii.ac.jp/GeoNLP/

of location names are coupled with their coordinates. TAC Knowledge Base Population (KBP) (McNamee and Dang, 2009; Ji et al., 2010a; Ji et al., 2010b) also has dealt with the task of entity disambiguation task as an entity linking problem where systems are required to link entity mentions to corresponding database entries. The difficulty of our task is that we have to detect fine-grained actual locations of location expressions which can include not only named entities but also expressions with general nouns (e.g. *the convenience store in front of the station*). Such disambiguation of general location expressions is a major issue since systems are required to predict actual entities from contextual information, etc. There is no previous work that addresses this kind of difficult task setting.

Also, how to collaborate with speech recognition systems is an important issue. Since the performance of the state-of-the-art speech recognition system is not perfect, we explore how to input emergency management information effectively and accurately from speeches.

Customizability is also an important requirement to make the system applicable for various local governments because processes of emergency management information can be different for each local government. We establish effective customizing methods through cooperating with local government employees.

### 2.3 Task 3: Environment Improvement for the Usage of Emergency Management Information Systems

In the Great East Japan Earthquake in 2011, it was indicated that emergency response requires information sharing between departments or organizations. In order to implement cross-organizational information sharing in disasters, it is essential to regularly hold emergency drills. Regarding drills, Hu et al. (2007) developed techniques for municipal employees to create drill scenarios reflecting local characteristics, by using samples. Motoya et al. (2009) examined emergency training management systems considering human resources development. The previous studies, however, were not focused on scenario contents which enable officials to enhance information sharing skills. It is essential in disasters to collect and handle information, create common operational pictures, and use them. Drill scenarios are required to check and

improve such skills. It is important to best utilize practical response data of the Earthquake to create such scenarios and implement drills to provide them.

For improvement of the environment for the usage of emergency management information systems, we prepare manuals that describe usage of the emergency management information systems and the database schemas used in the system. Also, we develop a scenario dataset for emergency drills. More precisely, we generate (a) hazard and damage maps for the purpose of training in emergency situations using a Web service we have developed. In addition, we create (b) a progress scenario of issues to be addressed based on the emergency management manuals, responses for flood damages, research results regarding cause-effect structures and the interview for employees in disaster affected local governments. Based on (a) and (b), we create a drill scenario dataset while appropriately including emergency responses for several emergency situations. Since we use the emergency management information database creation support system in training, we conduct situation annotation by using appropriate media consistent with the input interface of the system. Also, the drill scenario dataset will be used as a data for developing several structurization technologies in Task 2. To exploit the drill scenario data, we make the dataset capable, especially in size, for training machine learning models. This scenario dataset will also be used as an evaluation scenario for conducting demonstration experiments. In addition, we prepare a useful manual that describes the usage of the scenario dataset by clarifying required contents of the manual in the demonstration experiments.

### 2.3.1  Research Issues and Plan

In the improvement of the environment for the usage of emergency management information systems, we need emergency response data to develop drill scenarios. However, so far, digitized emergency responses have not been accumulated as archives. The records of emergency responses in Tohoku earthquake/tsunami have especially not been digitized, which can be effectively utilized to conduct emergency response training for potential Nankai Trough off the coast of western Japan. Developing training datasets and designing emergency management systems require actual emergency management records that are digitized and recorded in a unified way. However, there are several local governments which cannot provide such records. In order to address this issue, we develop a drill scenario data by enhancing limited emergency response data obtained from interviews for emergency responders by complementing them with damage information and status information. In addition, we take every possible means to develop drill scenario data by examining media used for situation annotation and considering status change due to emergency responses. After the development, we improve the data by asking employees in charge of disasters in local governments to check the data. By developing a multiple drill scenarios, we make the data applicable for various local governments.

To encourage broad use of the system, we demonstrate the system and conduct emergency response training in the local governments that cooperated with us. In the encouragement, we make the system cloud-based and construct a system which makes it easily available to conduct training in local governments through networks. In making the system cloud-based, we follow ISO22320, the standard of emergency management.

## 3  Current Status

We have already developed cooperative relations with some disaster affected local governments and have analyzed emergency response records provided by them in regards to how many contacts were received for each division, etc. From the analysis, we obtained fundamental information which should be considered in the development of database schemas. By analyzing the documents of the emergency management headquarter, we found that (i) updates of emergency information are mainly focused on 3 days after the disaster and (ii) understanding an overview of the disaster is easily accomplished by organizing and fixing items of emergency management information beforehand, etc.

Based on the data and the analysis, we developed a sample scenario data consisting of 100 entries and used it as reference data for designing the system. Table 1 shows some entries extracted from the developed scenario data.

Based on the data, we analyzed the task of creating emergency management information databases in the light of natural language processing, and designed a flow consisting of structur-

| Affairs division | Text |
|---|---|
| 本部:消火活動<br>Headquarters:Extinguishing | 17 時現在、15 時に発生した洞沢の火災は鎮火しました。<br>The fire occurred at 3 p.m. in Dozawa was extinguised. |
| 本部:物資調達<br><br>Headquaters:Procurement | ベビーおむつの在庫が支所にはありません。中上薬局から粉ミルクも確保する必要があります。<br>Diaper is now out-of-stock in the branch. It is also necessary to obtain the powder milk from Nakaue pharmacy. |
| 本部:交通規制対応<br>Headquarters:Traffic_regulation | 大谷の国道４５線は津波によって陥没しており、通行不能です。<br>Route 45 in Otani is currently not available because there is subsidence due to tsunami. |
| 避難所:開設・閉鎖<br><br>Shelter:Establishment/closing | 小泉中の避難所を副分団長が確認しました。昨日より開設とのこと。電気もガスも水も使えないので支援願います。<br>The vise-reader of the division checked a shelter in Koizumi junior high school. It has been established from yesterday. Utilities are not available. Support is requested. |
| 避難所:施設復旧<br>Shelter:Restoration | 津谷中学校は、停電していますが、教室を解放して避難者を受け入れています。<br>There is no electricity in Tsuya junior high school, but it is opening classrooms and accepting refugees. |
| 避難所:物資<br><br>Shelter:Goods | 寺谷からおにぎりが１００個の到着したので、仙翁寺にいる約２００人の消防団へ届けました。<br>We received 100 rice balls arrived from Teraya and sent them to a fire company which consists of about 200 members in Sennoji. |
| 避難所:仮設トイレ<br>Shelter:Temporary_lavatory | 小泉中学校です。至急、10 個ほど仮設トイレを設置できないでしょうか？<br>This is Koizumi junior high school. Please install about 10 temporary bathrooms as soon as possible. |
| とりまとめ様式・避難者<br>Summary:Refugees | 避難者の報告です。小泉中学校が約４００人、はまなすの丘が約３００人になります。<br>A report on refugees. There are about 400 refugees in Koizumi junior high school and about 300 refugees in Hamanasu-no Oka. |

Table 1: Samples of the developed entries of emergency management information.

ization of natural language texts, template selection, information extraction, and error correction by humans. In the structurization of natural language texts, we found that we need morphological analysis, chunking, dependency parsing, coordination analysis, entity extraction, event extraction, numerical expression recognition, location expression recognition, relation extraction, modality analysis, voice analysis, aspect analysis, existence analysis, clause relation analysis, discourse relation analysis and anaphora resolution.

We have started to develop linguistic analyzers required for the development of the proposed system, including a numerical expression recognizer, location expression recognizer, relation extractor, etc. As of now, the developed system is still in infant stage.

## 4 Conclusion

In this paper, we described an overview of our project, developing a system that assists structuring of emergency management information, and its current status. To utilize emergency management information in order to improve resilience against emergencies, we develop a framework that simplifies structuring emergency management information through various media or formats by exploiting natural language processing technolo-

gies. The tasks we address in this project are: (1) Designing standard DB schemas for emergency management information, (2) Developing a system that assists with creating emergency management information databases, and (3) Environmental improvement for the usage of emergency management.

In the future, we plan on establishing the database schemas commonly applicable for various local governments, and we progress with the development of the system that assists with creating databases with emergency management information using natural language processing technologies. After the development, we include the developed system into WebEOC, a standard emergency management information system, and conduct demonstration experiments with local governments to analyze technical and systematic issues for deploying the system into the field.

## References

Han Bo, Paul Cook, and Timothy Baldwin. 2012. Geolocation prediction in social media data by finding location indicative words. In *Proceedings of COLING 2012: Technical Papers*, pages 1045–1062.

Davide Buscaldi and Paolo Rosso. 2008a. Geowordnet: Automatic Georeferencing of wordnet. *Proc. LREC, Marrakech, Morocco*.

Davide Buscaldi and Paolo Rosso. 2008b. Map-based vs. knowledge-based toponym disambiguation. In *Proceedings of the 5th ACM Workshop On Geographic Information Retrieval (GIR 2008)*, pages 19–22.

Davide Buscaldi. 2010. *Toponym disambiguation in information retrieval*. Ph.D. thesis, .

Davide Buscaldi. 2011. Approaches to disambiguating toponyms. *SIGSPATIAL Special*, 3(2):16–19, July.

George R. Doddington, Alexis Mitchell, Mark A. Przybocki, Lance A. Ramshaw, Stephanie Strassel, and Ralph M. Weischedel. 2004. The Automatic Content Extraction (ACE) Program - Tasks, Data, and Evaluation. In *Proceedings of LREC 2004*.

Ralph Grishman and Beth Sundheim. 1996. Message understanding conference-6: A brief history. In *Proceedings of the 16th conference on Computational linguistics - Volume 1*, COLING '96, pages 466–471.

Mena B Habib and Maurice van Keulen. 2013. A Hybrid Approach for Robust Multilingual Toponym Extraction and Disambiguation. In *International Conference on Language Processing and Intelligent Information Systems, LP & IIS 2013*.

Mitsuhiro Higashida, Masahiro Sugiyama, Hideki Takeda, Tomomi Yamamoto, Yuji Maeda, and Haruo Hayashi. 2012. Analysis of information processing patterns appeared at emergency operation center training. *Outline of Social Safety and Science*, 30:93–96.

Zhexin Hu, Yasunori Hada, Toyoharu Itou, and Yashushi Saitou. 2007. A study on scenario making methods for disaster response exercised by local government personnel. *Journal of Social Safety and Science*, 9:271–278.

Heng Ji, Ralph Grishman, Hoa Trang Dang, Kira Griffitt, and Joe Ellis. 2010a. Overview of the TAC 2010 knowledge base population track.

Heng Ji, Ralph Grishman, Hoa Trang Dang, Kira Griffitt, and Joe Ellis. 2010b. Overview of the TAC 2011 knowledge base population track.

Kisung Lee, Raghu Ganti, Mudhakar Srivatsa, and Prasant Mohapatra. 2013. Spatio-Temporal Provenance: Identifying Location Information from Unstructured Text. In *IEEE Information Quality and Quality of. Service for Pervasive Computing (IQ2S)*.

Paul McNamee and Hoa Trang Dang. 2009. Overview of the TAC 2009 knowledge base population track.

Yutaka Motoya, Haruo Hayashi, Norio Maki, Keiko Tamura, Reo Kimura, and Kayoko Takemoto. 2009. Suggestions on how to design efficient training and management systems for personnel in charge of emergency responses, placing an emphasis on the process of developing human resources: A study based on the development and operation of the personnel training system designed for the cabinet office's division in charge of disaster prevention. *Journal of Social Safety and Science*, 11:203–213.