Combining Government and Linked Open Data in Emergency Management

Axel Schulz[1,2] and Heiko Paulheim[3]

[1] SAP Research
[2] Technische Universität Darmstadt
Telecooperation Group
axel.schulz@sap.com
[3] Technische Universität Darmstadt
Knowledge Engineering Group
paulheim@ke.tu-darmstadt.de

Abstract. In the Web 2.0 age, the challenges in emergency management have changed: in the past, it was a common problem that information was not available. Today, it is available in abundance, and what is needed is filtering and aggregating information to facilitate optimal situational awareness for emergency staff members. Our mashup *MICI*[1] shows how Open Government Data sources can be hooked up with Linked Open Data to provide useful dynamic views on emergency situations.

1 Motivation

For IT systems supporting an emergency staff, having the right information at the right time is a key requirement. Missing or incorrect information about fires, floodings, or power outages can increase the risk of wrong decisions with severe consequences. While in the pre-web age, the main difficulty was to collect the required information via phone or radio, it is nowadays available on the web in a timely manner – news channels provide up-to-date information, and citizens use Twitter or other social networks to share information about incidents. On the other hand, filtering and assessing that massive stream of information has become a key challenge in emergency management [5]. Without adequate strategies of handling the information overload, the situational picture can easily get cluttered, and emergencies can be misinterpreted when important aspects are overseen, which leads to an increased risk of wrong decisions.

Linked Open Data [2] is a collection of structured data, which can be used as background knowledge for adding assistance to staff members in emergencies. However, simply adding all possible information from Linked Open Data will not remedy those problems, as it will only increase the amount of data contained in a situational picture and make the information overload worse. Instead, background information must be added very carefully. The mashup *MICI* shows how background information from Linked Open Data can be used instead to assess,

[1] **M**ashup for **I**dentifying **C**ritical **I**nfrastructure

classify, and filter information gathered from the open government data [3], while only adding relevant bits of information from Linked Open Data. In a scenario using real, up-to-date fire brigade data from Seattle, we show how adding Linked Open Data can help classifying and structuring an otherwise unordered stream of data.

2 Prototype

Our prototype reads Open Government Data, e.g., as RSS data, which contains information about incidents. Users may define rules which classify the severity of an incident, e.g.: *if there is a fire within 50m radius of a gas station or a gas pipeline, the severity is high.* For evaluating such a rule, the background information about nearby objects, such as gas stations, is taken from Linked Geo Data [1], a data set within Linked Open Data which contains information about several objects with geo coordinates. For our prototype scenario, we use the fire call dataset from the city of Seattle[2], which provides a list of fire calls with type and coordinates, among others.

2.1 User Interface

The main screen of the prototype shows two views on the current situational picture, as depicted in Fig. 1. On the left-hand side, a list of incidents is displayed, which can be filtered according to the incidents' severeness. The right hand side shows a map view. When the user selects an incident in the list, that incident and all affected infrastructure objects are displayed in the map.

Rules are used both to find relevant infrastructure objects, as well as to calculate the severeness of an incident. To define which objects are relevant for which type of incident, users may create their own rules on the rule panel. To that end, they assign a set of object types to an incident type and define a radius and a degree of severity. For defining a rule, the user needs to know about the potential types of objects on which background knowledge is available, i.e., the ontology of Linked Geo Data defining object types such as gas stations, schools, etc. This ontology is used as a vocabulary for the user to define rules, while the vocabulary for the incident types depends on the RSS source used.

As the rule sets are relatively constant, editing the set of rules is something which is typically done once, and not at the time of an incident. During an emergency, the command staff will rather work with the main screen, which provides a clear situational picture to the emergency staff.

2.2 Architecture

Figure 2 shows the architecture of our prototype. The RSS reader reads incidents from an RSS source. The rule engine is responsible for classifying incidents based

[2] `http://data.seattle.gov/Public-Safety/Seattle-Real-Time-Fire-911-Calls/kzjm-xkqj`

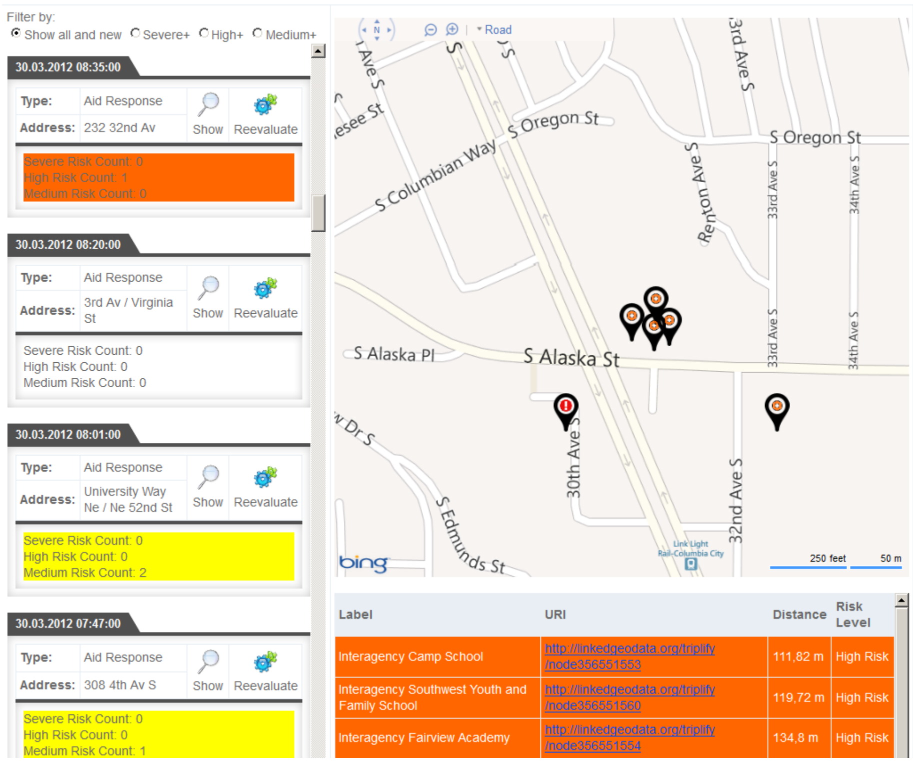

Fig. 1. Viewing incidents on the main screen. The list on the left hand side shows the current incidents, the map depicts a selected incident with identified critical infrastructure, such as schools.

on rules entered by the users, using background knowledge from Linked Geo Data. A rule is a tuple of the form:

$$\langle incident\_type, radius, object\_types, class \rangle \quad (1)$$

The incident type is provided by the RSS schema of the incident source, the object types are classes from Linked Geo Data, and the classifications are an ordered set of degrees of severity. The above example would be formalized as follows:

$$\langle Fire, 50, \{\texttt{lgdo:Fuel}, \texttt{lgdo:Pipeline}\}, severe \rangle \quad (2)$$

For classifying incidents, the rule engine uses the maximum of all radii defined for formulating a SPARQL[3] query to retrieve objects nearby the incident. Each rule is then evaluated against the set of retrieved objects, and the maximum classification of all firing rules is used (for example, if one rule states that an

[3] http://www.w3.org/TR/rdf-sparql-query/

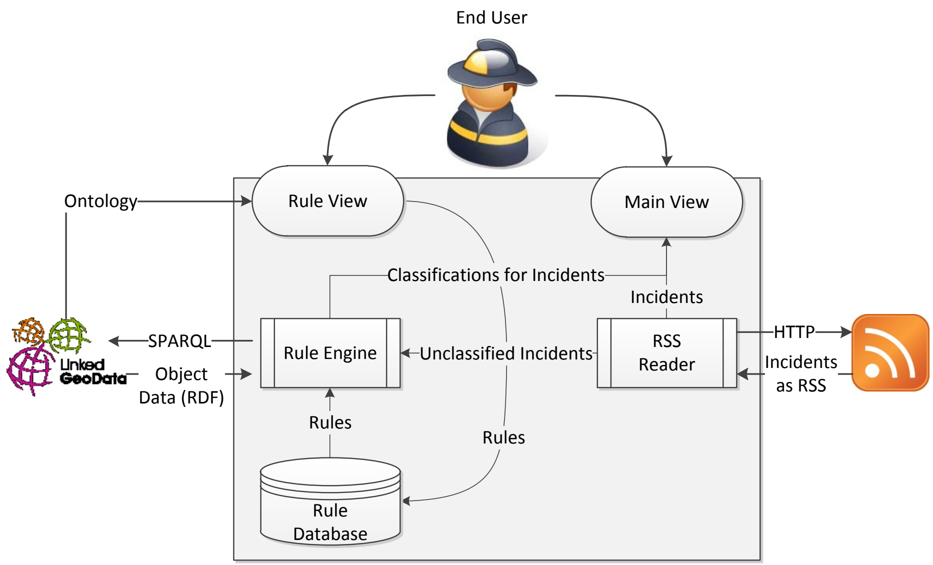

Fig. 2. Prototype architecture

incident is severe, while two more state that it is medium, the overall assessment is severe). Furthermore, those objects that made a rule fire are included in the incident information. Therefore, a rule serves two purposes: (1) to classify incidents and (2) to pick relevant information objects from Linked Open Data.

The processed incidents can then be used to provide intelligent views. For example, incidents may be filtered by severity, marked with different colors and/or symbols on a map, etc. By adding the corresponding objects responsible for making a rule fire, information on potentially harmed infrastructure (such as gas stations and pipelines) can also be provided to the end user, e.g., in a map view. Thus, the user may not only see that an incident is severe, but also get direct access to the objects that caused that rating, i.e., the name and the position of a specific gas station.

3 Conclusion and Future Work

With our mashup *MICI*[4], we have given a first glance on how Open Goverment Data data and information from Linked Open Data can be combined into a clear situational picture. Rather than simply adding all possible information from Linked Open Data to information items, which would lead to cluttering the situational picture, background information Linked Data is used to classify and filter given data, and only relevant pieces of information, such as critical infrastructure, are added to the situational picture. High-level rules defined which

[4] `http://mici.tk.informatik.tu-darmstadt.de/`

can be defined and understood by end users are facilitated both for classifying Open Government Data and to select relevant information from Linked Open Data.

For the future, we want to extend our system to allow for more Open Government and Web 2.0 data sources. A major challenge here is that not all data sources are already present in a structured form – some data sources, such as the popular Twitter messages, consist merely of short text messages. Turning such data into a structured data source for use in emergency management is a long term goal of our research [5].

While some data sources, such as the fire call data used in this scenario, already contain geotags, this is not true for all data sources. Therefore, automated approaches for predicting the geographic location of Web 2.0 data, such as Twitter messages, are highly desirable to make this approach more universal [4]. Using other Linked Data sets providing more detail information could improve the system as well. Furthermore, selectively adding more information from Linked Open Data (e.g., the phone number of a gas station) could increase the practical use of the mashup in real emergency use cases.

While creating the rules is fairly simple for end users, another approach would be to let experts classify example incidents, and induce the corresponding classification rules with rule learning algorithms. For the future, we envision a mixed approach where users can correct the automatic classification of an incident, and we use rule learning algorithms to refine the ruleset accordingly to cover the misclassified example. Further refining rules by additional attributes would also be interesting, e.g., a fire near a school may be more severe during day time (when it is likely that many people are in the school) than at night, while fires near gas stations are equally severe at all times.

Acknowledgements

The work presented in this paper has been partly funded by the German Federal Ministry for Education and Research (BMBF, 13N10712) "InfoStrom", and by the German Science Foundation (DFG, FU 580/2) "Towards a Synthesis of Local and Global Pattern Induction (GLocSyn)".

References

1. Auer, S., Lehmann, J., Hellmann, S.: LinkedGeoData – Adding a spatial Dimension to the Web of Data. In: International Semantic Web Conference (ISWC). (2009)
2. Bizer, C., Heath, T., Berners-Lee, T.: Linked Data - The Story So Far. International Journal on Semantic Web and Information Systems **5**(3) (2009) 1–22
3. Ding, L., DiFranzo, D., Graves, A., Michaelis, J., Li, X., McGuinness, D.L., Hendler, J.: Data-gov Wiki: Towards Linking Government Data. In: Spring Symposium of the AAAI. (2010)
4. Paradesi, S.: Geotagging Tweets Using Their Content. In: Twenty-Fourth International Florida Artificial Intelligence Research Society Conference. (2011) 355–356
5. Schulz, A., Paulheim, H., Probst, F.: Crisis Information Management in the Web 3.0 Age. In: 9th International ISCRAM Conference. (2012) to appear.