# Deployment of Semantic Analysis to Call Center

Takahiro Kawamura
and Shinichi Nagano
Corporate Research & Development Center,
Toshiba Corp.

Akihiko Ohsuga
Graduate School of Information Systems,
University of Electro-Communications,
Japan

## I. Introduction

In this paper, we present an application of text data triplification for a business. Since this is an in-company application from a laboratory to a division, we cannot describe it as “a success story of business-relevant, industrial deployments of Semantic Web technologies” in the CFP, although it will be useful as a case study.

Our company manufactures and sells consumer electronics ranging from refrigerators to TV sets, and it has recently been endeavoring to deal effectively with a number of inquiries about product malfunctions, which are gathered at a call center. Nowadays, moreover, if the response to an inquiry is mishandled, users tend to be complainers in some cases. A bad reputation then spreads widely via social media, that is, “flaming” occurs, and may greatly affect sales of all the company’s products. Making the response more problematic for operators at the call center is the difficulty of distinguishing whether the malfunction that is the subject of the inquiry is caused by a user’s way of using the product or a problem that accidentally occurs in an individual product, or caused by a problem common to the design or production phase of a particular model. In the case that an operator considers the malfunction to be the user’s fault at the initial stage, and it subsequently turns out to be the manufacturer’s fault, a firestorm may occur that may lead to lawsuits. The Consumer Affairs Agency in Japan and several law firms warn that the initial response to an inquiry is especially important in general. However, since pernicious complainers exist, if the manufacturer always considers the inquiry to be the manufacturer’s fault, the cost will soar.

Therefore, we proposed a method of comparing semantically analyzed social media information and the inquiry content. We triplify entries about product malfunctions on social media, and convert them to a network of Linked Data in advance. Then, by searching for the content of the inquiry to the call center in the network, we confirm whether the same issue is currently spreading on social media and whether the inquiry is the tip of an iceberg. If there is a similar entry on social media, it is determined whether the inquiry content is a malfunction common to a model and, if so, the operator offers a polite explanation to the user and a notification is sent to a quality control (QC) section. Moreover, if the entry has causal links connecting to users’ dissatisfaction and discontent, a notification with high priority will be sent to the quality control section.

We, that is, our laboratory, brought the above-mentioned advantages to the attention of a division of our company, which manufactures and sells consumer electronics, and then received a research contract with a certain amount of R&D expenses[1].

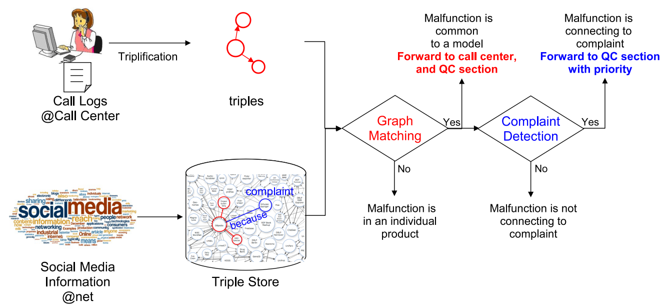


Fig. 1. Search flow of inquiry contents from Linked Data

## II. Triplification of Social Media Information

To create a training dataset, firstly, we divided each sentence in the dataset into chunks of semantically consistent words by using Part of Speech (POS) analysis and syntactic analysis, and then manually labeled one of eight properties, namely, Subject, Action, Object, Location, Time, Modifier, Because, and Other, to each block. We then used conditional random fields (CRF) as a learning model, which is an undirected graphical model for predicting a label sequence for a sequence. The key point of the proposed method is that we also constructed approximately 250 annotation rules using the result of syntactic analysis and the predefined ontology, for example, a noun before a postpositional particle ‘WO’ corresponds to OBJECT in a Japanese sentence, and a sentence after a word ‘NAZANARA’ (because) and a sentence before the word have a causal relation, and so forth. We then decided which of the CRF estimation and the rule decision should be adopted based on the estimation probability of CRF.

In addition, we determined identities of values (chunks), that is, entity linking, so that values of Subject, Object, etc. that have the same meaning refer to an identical node in the network, as much as possible. Finally, we unified the values that are determined to be identical to a node whose label is a typical value.

## III. Matching between Inquiry Content and Linked Data

Figure 1 presents the flow when an inquiry is received at a call center. When the call center receives an inquiry from a user, an operator records the summary of the inquiry content as two or three sentences (call log). Each sentence is

[1] approx. ten million yen for a half year

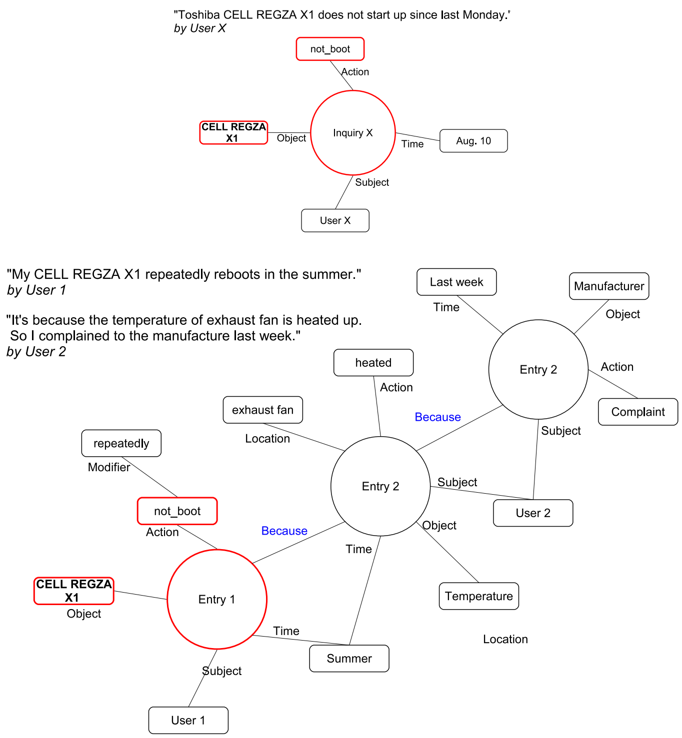


Fig. 2. Linked Data graph for an inquiry content (above) and correspoinding social media information (below)

triplified in the format of $< S_i, V_i, O_i >$, and then triples that have the same structure as the sentence are searched in the triple store. As a result, if a triple with the same structure as the inquiry content is found, we determine that the problem does not concern an individual product, but is common to a model. Moreover, the number of triples with the same structure is regarded as an amount of topics on social media. When querying the triple store to find $S_s, V_s, O_s$, we also use the method of entity linking described in II. Example graphs of social media entries and inquiry content are shown in Fig. 2, where each sentence has a sentence ID node and at most eight properties.

## IV. Experiments on Triplification and Matching

### A. Triplification of social media

In an experiment, we collected entries about a TV set manufactured and sold by our company from a well-known review site in Japan[2], and then conducted labeling, learning, and estimation with the method described in the previous section. The dataset is 197 sentences for three months, and evaluated with 10-fold cross-validation. Table I shows the combined result of the CRF estimation in the case of the estimation probability $p > 0.6$ or the rule decision, otherwise.

TABLE I. Extraction accuracy for each property

| (%) | SUB. | OBJ. | ACT. | LOC. | TIME | MOD. | BCOZ | Ave. |
|---|---|---|---|---|---|---|---|---|
| Precision | 85.7 | 88.8 | 96.9 | 63.6 | 100.0 | 88.2 | 100.0 | 94.1 |
| Recall | 100.0 | 92.7 | 95.4 | 46.7 | 67.9 | 91.3 | 100.0 | 94.1 |

Weighted Average (Ave.), which is an average value according to the number of each property, indicates that the combined method we proposed achieved accuracy of 94.1%.

[2]http://kakaku.com

The accuracy of the Location property is lower than that of other properties because of the shortage of geographical names registered in the system. The low accuracy of the Time property seems to be attributable to the difficulty of distinguishing it from the Modifier property. We also confirmed that extraction of the causal relation is feasible, since the accuracy of the Because property is high.

The division to which we provided this result commented that the 94.1% extraction accuracy is satisfactory, but pointed out that on this occasion social media information was collected for a certain period and converted to a graph (Linked Data), and therefore the graph represents a snapshot. Opinions expressed on social media are continually changing from product release to malfunction discovery and manufacturers' responses, and thus such time-series variations should be represented in the graph. In addition, users' complaints are of varying strength, and thus they should be divided into multiple stages from a weak complaint to a strong complaint. Therefore, we intend to prepare more detailed properties for representing various nuances of verbs.

### B. Matching between inquiry content and social media

In the experiment, we first extracted 220 call logs (summaries of inquiry contents described by operators) from 25,459 logs about our company's TV sets for a month, September 2012. We then compared them with social media information that was triplified as described in IV-A. Finally, the matching results between the call log and part of social media were manually checked, and then the accuracy of the matching was calculated. The result is shown in Table II.

TABLE II. Matching accuracy of inquiry contents (Ave.)

| No match | | Match | |
|---|---|---|---|
| No data | Triplification Error | Precision | Recall |
| 9.1% | 13.6% | 88.2% | 33.3% |

The fact that the precision of call logs to social media graph was about 90% indicates that checking the same entry on social media as a call log is possible. Since the recall was low, however, we found that it is difficult to deduce how widely the call log is spreading on social media from this result. The recall was low because there are several expressions that represent the same condition and content on social media, and also the method of entity linking mentioned in II is insufficient to unify them.

The division to which we provided this result commented that when an inquiry is received at a call center, it is not possible due to time constraint that an operator performs keyword search with appropriate keyword expansion, and find the same entry as the inquiry content on social media, but this system automated comparison between call logs and social media using semantic search with word identification and word relation. The comment also indicated that in future when the malfunction of a model is spreading on social media, an alert should be transmitted before receiving the call log.

## V. Conclusions and Future Work

Future works include performance evaluations. We have developed the system and are in the trial phase. In the future, we intend to identify issues that may arise through the actual operation of the system, and further improve the system.