Automatic Question Generation to Determine Roles During a Crisis

Marten Teitsma
University of Amsterdam
The Netherlands
m.teitsma@uva.nl

Jacobijn A.C.Sandberg
University of Amsterdam
The Netherlands
j.a.c.sandberg@uva.nl

Marinus Maris
University of Amsterdam
The Netherlands
m.maris@uva.nl

Bob J.Wielinga
University of Amsterdam
The Netherlands
b.j.wielinga@uva.nl

***Abstract*—Traditional information systems for crisis response and management are centralized systems with a rigid hierarchical structure. Here we propose a decentralized system, which allows citizens to play a significant role as information source and/or as helpers during the initial stages of a crisis. In our approach different roles are assigned to citizens. To be able to designate the different roles automatically our system needs to generate appropriate questions. On the basis of information theory and a restricted role ontology we formalized the process of question generation. Three consecutive experiments were conducted with human users to evaluate to what extent the questioning process resulted in appropriate role determination. The result showed that the mental model of human users does not always comply with the formal model underpinning the questions generation process.**

***Keywords*-Crisis Management, Ontology, Human-Centered Sensing, Theory of Strongly Semantic Information, Situation Theory**

I. Introduction

When disaster strikes, information gathering is of great importance. During the response phase, when the disaster has just happened, information is most needed but also most scarce. It is during this phase that people and emergency services plan actions in an information twilight. In this paper we describe a formal method to support automatic question generation in an efficient way. This process aims to determine, which roles people can play and how they can help with an adequate response to the disaster. Several experiments with human users were conducted to validate the question generation process and the role determination this results in.

Information technology can be of use to gather information during the so called "golden hour" (i.e., the first sixty minutes after a severe trauma) [11]. But when it comes to information gathering, a focus on a centralized approach has been the usual course [5]. A centralized structure comes along with a strong hierarchical reporting structure, which has been the model for use by the emergency services. Such systems tend to ignore the public as a source of information. Our intended system is (partly) decentralized, i.e., the application runs on a mobile phone, and makes use of ordinary people who happen to be in the disaster area. Until now grassroots participation of citizens during a disaster as a valuable contribution to information gathering has not been fully appreciated by emergency services and other formally involved parties [9]. Due to this lack of appreciation, efforts to develop a technological platform to enable such participation are limited. It has been found however that, even during the most agonizing moments, people tend to help each other and can act rationally [2].

Making use of humans to gather information is the central subject in the new emerging field of Human-Centered Sensing (HCS) [6]. The here proposed application is typified as a participatory sensor because humans are producing information and not just facilitating the gathering of data as in opportunistic sensing e.g., a mobile device recording background noise. By answering questions the human observers can help, making clear what the situation is.

In the context of disasters it is important to be aware of the short time span available. Our assumption is that people do not want to be engaged in a time consuming questionnaire when all around them the world turns upside down. Therefore we designed a very simple ontology, which leads to a limited number of questions. This formalization is needed to automate the question generation process. A non-formalized communication would engage too many people in a call or operation center.

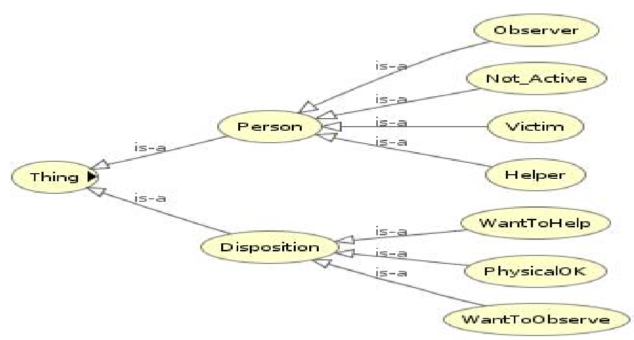

Figure 1. An ontology for roles during a disaster

An example: suppose a hurricane is expected to hit a large urban area. The people in the area already have our application installed on their mobile phone, which guides them through the querying process. After the initial phase of the disaster the users are asked a couple of questions to determine their physical condition as well as their need for help and their inclination to help others and their willingness to observe. This information is used in the role determination process. Part of future research will be that when people are classified in different roles, they are asked to perform specific tasks. A *Helper* for example gets the task to go to a place where she can find a *Victim* who is in need of help. Or, a *Victim* is asked to describe the injury he has. Such information will be helpfull to the *Helper* when she is helping this *Victim*. Furthermore, an *Observer* may be asked to give information about his surroundings and tell about the number of people he sees who are hurt.

In this paper, we examine and test the way questions to determine roles can be automatically generated. First, we will provide the theoretical background by discussing some related work. Next, we propose a role determining question generator based on an ontology. Several experiments, conducted to validate the question generation process, are presented and discussed.

II. Generating questions

Our application strives to determine the role people can have during the response phase of a disaster. Herefore we first have determined, which roles we can discern and define. The definition of roles is done in an ontology. Each concept of a role has certain properties. To avoid a combinatorial explosion we took dependencies into account. These dependencies result in a number of impossibilities, which then can be ruled out in the determination process. To determine which question to ask first, we have developed a specific method, i.e., semantic strengthening based on the Theory of Strongly Semantic Information [4].

A. Ontology

An ontology is a set of concepts and their interrelations, which formally represents objects in a particular domain. Due to the formalism it is possible to reason about the concepts and their properties. To design an ontology we used Protégé-OWL [7]. The semantics of OWL is founded on Description Logic, which is a decidable but still expressive formalism [1].

Because we use properties to generate questions the number of properties per role must be kept to a minimum. Furthermore, they shouldn't be ambiguous. The third requirement for the properties is that they must be maximally subjective, i.e., the answer must rely on the thoughts and feelings of the person herself. Whether people want to help other people or not depends on their disposition to help. The same subjective perspective must be applied to the willingness to observe and even the physical condition of the people who we approach.

After a disaster has struck it is important to quickly distinguish between (groups of) active and non-active people. The non-active people can be victims who are affected by the disaster in such a way that they need help and people who are not physically affected but for some reason don't want to be active. The active people are helping to mitigate the effects of the disaster. They do this by directly helping other people or by observing and generating information useful for emergency services or the mentioned helpers. And thus consists our classification of the roles: *Victim*, *Helper*, *Observer* and *Not-Active* (see also Fig. 1).

Our ontology consists of definitions of the form:

$$\begin{aligned} Observer \equiv Person \\ \cap \exists hasDisposition(PhysicallyOK) \\ \cap \exists hasDisposition(wantToObserve) \end{aligned} \tag{1}$$

which says that *Observer* is equivalent to being a member of the set *Person*, which has the restriction of being member of the two sets of being physically OK and wanting to observe. In the ontology, other concepts like *Gender*, *Age* and Location are also described. These are concepts we want to use in the development of our system where we also use more personnel characteristics.

B. Dependencies

To discuss the information we need and the combination of different pieces of information we use the terminology developed in Situation Theory by Devlin in [3]. In Situation Theory a piece of information is called an infon, which is formally described as a tuple of the form:

$$\langle\langle R, a_1, ..., a_n, 0/1\rangle\rangle \tag{2}$$

where R is a n-place relation, and $a_1, ..., a_n$ are variables representing objects appropriate for R. The last item is the polarity of the infon. When it is "1" the infon is true given a particular situation, otherwise false and "0". We depict a situation as a defined set of infons. This is the minimum number of facts defining the situation.

Trying to determine which situation is the actual situation, one easily creates an enormous amount of possible situations. The number of answers to a question determines how many situations are possible as description of the real situation. A "yes" or "no" as answer gives per question two possible situations and the addition of "I do not know". results in three possible situations. When having more than one question this easily leads to great numbers of possible situations. For example, 4 questions with each 3 possible answers gives 81 possible situations. One has to constrain this combinatorial explosion. In the previous section we discerned four different roles based on four different properties. Each property is a piece of information we want to ask about. Such a property will be formulated as follows:

$$\langle\langle \text{hasDisposition}, wantToObserve, p, t, l, 1\rangle\rangle \tag{3}$$

where p, t and l are parameters for a specific person, time and location. Taken together, such infons can describe a situation of a person. And so having four properties gives

 ISBN: 978-1-61208-163-2

	σ_1	σ_2	σ_3	σ_4
S_1	0	0	0	0
S_2	0	0	0	1
S_3	0	0	1	0
S_4	0	0	1	1
S_5	0	1	0	0
S_6	0	1	0	1
S_7	0	1	1	0
S_8	0	1	1	1
S_9	1	0	0	0
S_{10}	1	0	0	1
S_{11}	1	0	1	0
S_{12}	1	0	1	1
S_{13}	1	1	0	0
S_{14}	1	1	0	1
S_{15}	1	1	1	0
S_{16}	1	1	1	1

Table I
TABLE WITH POSSIBLE SITUATIONS WHEN HAVING FOUR INFONS

16 (2^4) possible situations as you can see in Table I. Here S_{15} describes an *Observer* when σ_1 is the infon, which says someone is a *Person*, σ_2 describes that someone is *PhysicallyOK* and σ_3 that this person *wantToObserve*. We then restrict the number of possibilities by determining dependencies between the properties.

There are three dependency relations in our ontology: the relation between "being physically OK" and "wanting to observe" and the relation between "wanting to observe" and "wanting to help". Because of transitivity we can detect a third dependency between "being physically OK" and "wanting to help".

This definition of concepts results in sets, which are subsets of other sets:

$$WantingtoHelp \subseteq WantingtoObserve \subseteq PhysicallyOK \subseteq Person \quad (4)$$

This equation says that the set of people who want to help is a subset of the people who want to observe, which is a subset of the people who are physically OK, which is a subset of persons. Here we see that when someone being physically OK implies being a person. And when someone wants to observe it is implied he is physically OK.

Knowing the dependencies in the system would make it the most efficient strategy to ask after whether people want to observe. But then, we suppose these people know that answering "yes" means they want to observe *and* are physically OK, which is a supposition we can not make. In a system with logical dependencies, one should not expect that all the varieties given in Table I do have an even chance of becoming real. It may even be so that some situations are impossible as outcome of a deliberation. The dependencies we formulated determine that situations in our system are possible or impossible. Whether a situation is possible or impossible is not known to the users of the system. Because we know there is a difference between the logic of our system and the mental model of the user, our system has to restrict the situations to possible situations and rule out the impossible ones. How we keep users away from these impossible outcomes is shown in the next section. First the impossible situations have to be determined.

The dependencies we have defined in the ontology restrict all the situations as mentioned in Table I to possible situations. Because all the roles are dependent on σ_1 this infon must necessarily be part of the situation. Looking at Table I, it is obvious which situations are impossible: $S_1 ... S_8$. But also S_{10}, S_{11} and S_{12} are impossible, because in these situations people want to observe or help but are not physically OK. At last, S_{14} is impossible because this person wants to help but not observe, which we also ruled out as possible.

C. Semantic strengthening

Now we know which situations are possible, we can determine after which infon we have to ask first. What we are after is an order of questioning, which leads to the roles as defined in the ontology. The roles are defined by their properties, which are represented as infons in the situations. Dependencies result in restricting the possible situations and excluding the impossible ones. But these restrictions are not known by the persons who use our system. In this section we describe a method to preclude the impossible situations or prohibit the assignment of roles not in line with our definition of these roles.

The order of questions can be found by using a method familiar to semantic weakening as described in [4]. With semantic weakening a series from total vacuity to a minimum vacuity is created. A statement has a minimum vacuity when it refers to the minimum number of situations. Total vacuity for a statement corresponds to a tautology in a specific domain because it is always true. Decreasing the number of situations, which are compatible with the true situation, increases the quantity of informativeness. Semantic weakening is done by connecting the infons, which constitute the situation by more and less disjunctions instead of conjunctions. The number of supported situations divided by the total number of possible situations is called the degree of vacuity. When, in the context of a probability experiment, which resulted in Table I, we make the statement $\sigma_1 \wedge \sigma_2 \wedge (\sigma_3 \vee \sigma_4)$, the situations S_{14}, S_{15} and S_{16} support the statement. The situation S_{13} is not supported because σ_3 and σ_4 are both false in this situation and $\sigma_3 \vee \sigma_4$ does not result in a true statement. Two disjunctions results in the (compound) infon $\sigma_1 \wedge (\sigma_2 \vee \sigma_3 \vee \sigma_4)$. This infon complies with even more situations: first off course S_{14}, S_{15} and S_{16}, and then also with S_{10}, S_{11}, S_{12} and S_{13}. When making the statement $\sigma_1 \vee \sigma_2 \vee \sigma_3 \vee \sigma_4$ all but S_1 is supported.

The method we use, *semantic strengthening*, is keeping the truthfullness when bypassing impossible situations. In

Number	Hypothetical role	Scenario
1	Victim	During the earthquake you were just drinking coffee in the kitchen. When you noticed the first trembles your ran out of the house but unfortunately a lot of debris was falling down and hit you. You have broken your leg and are not able to move. The telephone rings.
2	Not-Active	You woke up in the middle of night when a police car was riding down the street calling everybody out of bed and warning for an immediate flooding. The police warned not to flee but instead look for a high place and take food and drinks with you. You immediately went to the refrigerator took food and drinks and climbed through the bedroom window to the roof. But now you are sitting there and it is getting colder and darker. The streetlights are not burning anymore, probably because the power is down and you hear water streaming but see nothing. You are getting afraid and what is even worse you lost your glasses so you can't see very clear. After a while the telephone rings.
3	Observer	After the first trembles you and your family ran out of your house. Luckily everybody came out of the house and now you are on the street. Your youngest child is only 3 months old and is sleeping now in your arms. Your 4 year old son is very excited and very wild probably because he is afraid. Your wife has quite a job to handle him. Your house has big cracks in it and you are afraid to go inside. Then the telephone rings.
4	Helper	During the earthquake you were walking in the park with your dog. You saw houses collapse and after five minutes when the earthquake seemed have come to an end you went for your house. But your house wasn't standing any more and collapsed like most of the houses in the street. Now you are in the street and the telephone rings.

Table II
A PART OF THE SCENARIOS FOR THE EXPERIMENTS

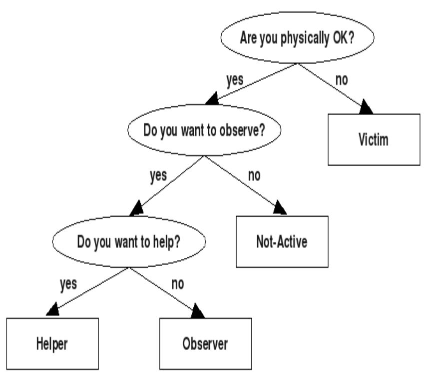

Figure 2. A question tree

our method we place emphasis not on the disjunctions but on the conjunctions. And the conjunctions are placed in such a way that there is no loss of truthfullness and impossibilities are ruled out.

The efficiency of the order of questioning is maximal, i.e., after each answer the total number of situations, as given by Table I is cut in half. It is important to be aware of the order by which the questions are asked. The specific order precludes the impossible situations as an outcome of this questionnaire. With our specific ontology this would result in a question tree as shown in Fig. 2.

III. EXPERIMENTS AND RESULTS

We conducted three experiments to investigate whether the questions we ask to determine the role of the user are indeed self explanatory and lead to appropriate role determination. Different disasters like an earthquake, flooding or a bombing were used to describe a situation where people are involved in, immediately after the occurrence. For each scenario a hypothetical role was envisaged i.e., the specific role, which was implied by the ontology should follow from the scenario. The goal of the experiments was to find out whether human participants answered the questions posed in the same way as hypothesized by our theoretical framework. Examples of the scenarios can be found in Table II.

A. Analysis

For the analysis of the data four measures were computed: the Matthews Correlation Coefficient (MCC) for correlation [8] and the F_1-score for accuracy [12], recall and precision. The MCC (also known as the ϕ-coefficient) is a measure of correlation between what is actual and what is predicted by a system or humans as in this case. Therefore so-called confusion matrices were needed to compute the measures. First is explained how we constructed the confusion matrices, followed by an elaboration on the measures and then the experiments are discussed.

As described in section II, the answers to the questions were used to compute the determination of a role. In the ontology, four roles were defined. To analyse the results as shown in Table III we constructed for each experiment four confusion matrices. An example may be helpfull. Of the four roles each scenario shown to the participants had an expected or actual role, which was envisaged e.g., *Victim*.

 ISBN: 978-1-61208-163-2

When the participant answered the questions so that the result was that he was a *Victim*, this is marked as “true positive” in *this* confusion matrix. When the participant was determined as being a *Not-Active*, *Observer* or *Helper*, this is marked as “false negative”. When another scenario was presented, with another envisaged role e.g., *Helper*, and the participant was determined as *Victim*, this is marked as “false positive” in *this* confusion matrix. When the participant was determined in that scenario as something other than *Victim*, this is marked as “true negative”.

We use four measures to interprete the results. MCC is used to tell whether there is a correlation between the actual and predicted values. It is a robust coefficient because it does not deviate when classes of different size are considered. MCC variates between -1 and +1 where -1 indicates a perfect negative correlation and +1 a perfect positive correlation, 0 indicates a random relation. The F_1-score is a measure of accuracy and varies between 0 and 1 where 0 indicates no accuracy at all and 1 a perfect accuracy. The F_1-score is the harmonic mean of the recall and precision. The recall (also called sensitivity or true positive rate) is a measure of how many of the actual situations are determined as such. Precision gives a measure of how many of the predicted situations are actually these situations.

Forty students participated in the first experiment, all of them male and between the age of 18 and 22. Eight scenarios, not very different from the four shown in Table II, were constructed in the english language. Each role was represented twice. The participants were asked to read four of the eight scenarios. These four always represented all four possible roles. As instruction, the participants were told to imagine being in the situation described by the scenario. Each scenario ended with the announcement that the telephone rings and then the participant answered the questions that were subsequently posed in Fig. 2.

The results of the first experiment are summarized in Table III. In this table one can see that actual values were most predicted when the participants were confronted with the *Victim* and *Helper* scenarios. And it shows a bias to the role of *Helper* when reacting on the *Not-Active* and *Observer* scenarios.

When analyzing these figures as in Table IV a very low value for correlation is measured except for the *Victim* scenarios. For the *Victim* scenarios the accuracy is relative high. For the *Not-Active* scenario the correlation is even negative, i.e., it has a reverse correlation. For *Observer* and *Helper* the correlation has a low value. For *Helper* this is a consequence of the high value of “false positive“ in the confusion matrix, which is also reflected in the low value for ”precision“. We then combined the roles of *Victim* and *Not-Active* and *Observer* and *Helper*. The correlation is still low and for *Victim* even declining. But for all other scenarios the correlation is improving. The same can be said of the accuracy.

Experiment 1	Predicted value			
Actual role	Victim	Not-Active	Observer	Helper
Victim	23	4	6	7
Not-Active	2	3	5	30
Observer	2	3	14	21
Helper	1	6	6	27
Experiment 2	**Predicted value**			
Actual value	Victim	Not-Active	Observer	Helper
Victim	47	1	2	9
Not-Active	11	8	6	34
Observer	3	4	14	38
Helper	5	5	5	44
Experiment 3	**Predicted value**			
Actual value	Victim	Not-Active	Observer	Helper
Victim	37	0	0	1
Not-Active	1	10	6	21
Observer	1	1	14	22
Helper	0	1	6	31

Table III
RESULTS OF EXPERIMENTS

Experiment 1	MCC	F_1	Recall	Precision
Victim	0,61	0,68	0,58	0,82
Not-Active	-0,05	0,11	0,08	0,19
Observer	0,23	0,39	0,35	0,45
Helper	0,17	0,43	0,68	0,32
Passive	0,28	0,52	0,4	0,73
Active	0,28	0,69	0,85	0,59
Experiment 2	**MCC**	**F_1**	**Recall**	**Precision**
Victim	0,67	0,75	0,80	0,71
Not-Active	-0,23	0,21	0,14	0,44
Observer	-0,17	0,33	0,24	0,52
Helper	-0,07	0,48	0,75	0,35
Passive	0,44	0,66	0,56	0,79
Active	0,44	0,75	0,86	0,67
Experiment 3	**MCC**	**F_1**	**Recall**	**Precision**
Victim	0,95	0,96	0,97	0,95
Not-Active	0,39	0,40	0,26	0,83
Observer	0,26	0,42	0,37	0,48
Helper	0,30	0,51	0,74	0,39
Passive	0,63	0,76	0,63	0,94
Active	0,63	0,83	0,96	0,72

Table IV
MCC, F_1, RECALL AND PRECISION FOR THE EXPERIMENTS

Because the first experiment was done with a very homogeneous group of young men we did the second experiment with a more heterogeneous group. Of this group 15.25% was woman and 33.9% of all the participants older than 22 year. In this experiment we also made the scenarios more explicit. Four of these scenarios can be found in Table II. Furthermore, we used a flow diagram per scenario to collect the answers for that scenario. In this experiment the scenarios were read in two groups: the first group read the scenarios 1-4 and the second group read the scenarios 5-8.

The results can be found in Table III. Although the raw results look a lot like those in experiment 1, i.e., the actual role was most predicted for *Victim* and *Helper* and a bias towards

 ISBN: 978-1-61208-163-2

the role of *Helper*, the analysis is very different as shown in Table IV. The scenario for *Victim* has a relative high value for correlation as in experiment 1 but the other scenarios score a negative value for correlation. When combining the roles as in experiment 1 this negative correlation reverses to a higher correlation than in experiment 1. The number of $0,44$ for MCC is still not high and should be considered ”positive“ but not ”strong positive“. The accuracy is also improving as are recall and precision.

In the third experiment 38 students participated, all of them male and between the age of 18 and 22. The third experiment was conducted with a different instruction and a different language. This experiment was in Dutch, which is the native language of most of the people we did the experiment with. We introduced the questions beforehand and gave one example of the dependencies we had defined. The scenarios were the same as in the second experiment (see Table II) but translated of course.

The results can be seen in Table III. As before the actual role was most predicted for *Victim* and *Helper* and the bias towards *Helper* can be seen. In Table IV figures of the MCC F_1, recall and precision are given. As can be seen there is a positive correlation for all the roles and for *Victim* even a very strong correlation and accuracy. When the roles are combined as before this correlation gets stronger for all the roles except for *Victim*. Moreover, the improving of the correlation and accuracy shown in experiment 2 continues.

IV. Discussion and conclusion

Each successive experiment showed an increased correlation between the actual role described in a scenario and the predicted one, which the participants selected after answering the questions. This is shown in Table III, where the predicted role in each column has the highest number of predictions in the third experiment.

As could be expected, adding a flow diagram, using native language and giving an adequate introduction is important for understanding the concepts we use for questioning. Furthermore, we can conclude that there is a difference between the formal definition of the concepts in the ontology and the semantic interpretation people have of these concepts. Morever, the meaning of concepts can, as we have seen, not only vary among people but also between people and systems. This discrepancy is shown in this experiment by different choices people make in answering the questions some of which were formally ruled out by our system. People do not straightforward comply to formal reasoning. This difference is even greater when refering to concepts denoting subjective situations, which intentions such as ”the willingness to help“ are. Hence, for the sake of disambiguation between such situations, the reasoning that the system does on the basis of the answers of people, ought to be augmented by verifying and confirming the answers provided.

Further research will be done to develop a model of commonsense reasoning in the context of enhancing Situation Awareness. Such a model will consist of basic concepts, which are information-rich and common in use [10]. The system we use will be a ”hybrid model”, which uses formalized methods to generate questions while incorporating possible mental models.

References

[1] F. Baader and W. Nutt. Basic description logics. In F. Baader, D. Calvanese, D. L. McGuiness, D. Nardi, and P. F. Patel-Schneider, editors, *The Description Logic Handbook. Theory, Implementation and Applications*, pages 45–104. CUP, Cambridge, UK, 2007.

[2] L. Clarke. Panic: Myth or reality? *Contexts*, 1(3):21–26, Fall 2002.

[3] K. Devlin. *Logic and Information*. Cambridge University Press, Cambridge, UK, 1991.

[4] L. Floridi. Outline of a theory of strongly semantic information. *Minds and Machines*, 14(2):197–221, 2004.

[5] N. Goodman and R. Langhelm. Passive disaster reporting through mobile social networking technology. In F. Fiedrich and B. Van de Walle, editors, *Proceedings of the 5th International ISCRAM Conference*. ISCRAM, May 2008.

[6] M. Jiang and W. McGill. Participatory Risk Management: Managing Community Risk Through Games. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, pages 25–32. IEEE.

[7] H. Knublauch, M. Horridge, M. Musen, A. Rector, R. Stevens, N. Drummond, P. Lord, N. Noy, J. Seidenberg, and H. Wang. The protégé owl experience. In *Fourth International Semantic Web Conference (ISWC2005)*, Galway, Ireland, 2005.

[8] B. Matthews. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, 405(2):442–451, 1975.

[9] L. Palen, S. R. Hiltz, and S. B. Liu. Online forums supporting grassroots participation in emergency preparedness and response. *Commun. ACM*, 50(3):54–58, 2007.

[10] E. Rosch, C. Mervis, W. Gray, D. Johnson, and P. Boyes-Braem. Basic objects in natural categories. *Cognitive psychology*, 8(3):382–439, 1976.

[11] J. van de Ven(ed.). The combined systems project. Technical report, DECIS, Delft, The Netherlands, 2006.

[12] C. van Rijsbergen. Foundation of evaluation. *Journal of Documentation*, 30(4):365–373, December 1974.

 ISBN: 978-1-61208-163-2