Case-Based Retrieval and Adaptation of Regulatory Documents and their Context

Andreas Korger[1] and Joachim Baumeister[2,3]

[1]Angesagt GmbH, Dettelbachergasse 2, D-97070 Würzburg
[2]denkbares GmbH, Friedrich-Bergius-Ring 15, D-97076 Würzburg
[3]University of Würzburg, Am Hubland, D-97074 Würzburg

Abstract. Regulatory documents are required or provided by authorities in many domains. They commonly point out relevant incidents for specific scenarios. For those they have to present suitable preventive and reactive measures. We introduce an approach to connect a case-based description of the incidents structure with a case-based description of the according context. This paper shows how to use case-based methods to retrieve, adapt, and reuse incidents descriptions. Subsequently they are used to generate new regulatory documents via case-based reasoning.
Case-based reasoning Experience Management Knowledge Management SKOS Semantic Relatedness Natural Language Generation.

1 Introduction

A *regulatory document* describes incidents that are likely to happen in a certain situation. Preventive measures are elaborated to avoid the occurrence of relevant incidents and adequate reactions are proposed. Further, harmful consequences are to be avoided or mitigated. This underlying structure is represented in the documents structure. Popular examples of regulatory documents are public events, for the handling of hazardous material or industrial workplace safety. For a festival, a regulatory document would describe incidents such as fire and relevant measures like the allocation of fire-extinguishers. The overall goal of this work is to support domain experts in writing regulatory documents. Fundamental considerations have been presented in preceding works [9, 10]. For instance methods for documentary adaptation using a combination of ontological document description and case-based reasoning. We extend ontologies to represent these special parts of documents containing regulatory information.

We use natural language processing to connect graph-based and textual knowledge representation. Our goal is to retrieve and adapt passages of documents depicting such regulatory knowledge for usage in another context.

For generating a new document, we need to answer three questions. Which incidents are likely to happen? Which preventive and reactive measures are suitable for each incident under a certain context? How important is each measure? The last question pays attention to the fact, that a limited budget and time does not allow for the implementation of all preventive measures. This sums up to consider a convenient context-based ranking for the incidents and measures. The presented approach is a general framework and easily adaptable to domains providing textual as well as ontological information for the context dependent classification, prevention of, and reaction to incidents.

For reasons of simplification and consistency we give examples of the domain of public events. Our approach is driven by theoretical and case-based considerations we describe in the first part of this work. Then we present the experimental setup we used to install a case-study showing practical capabilities of our approach. We finish with related work as well as with discussions and future work.

2 Ontology Model for Incident Assessment

We make the assumption that there exists a corpus of regulatory documents of a certain domain. The documents are sub-classified into passages, that are connected with incidents or the according measures. Those passages of text are called *information units*. The work of identifying these passages was done by domain experts. All information beyond the textual corpus of available documents is coded into a knowledge base [3]. This knowledge base consists of *entities* (ontological concepts) as logical units and the *relations* between them. An entity may be for instance an action, an agent, an event, or a resource. In a textual context an entity is coded or described as one word (term) or more words up to some sentences. Entities may be composed of sub parts in an arbitrary manner. In the following, we introduce the basic concepts of our scenario.

Definition 1. *Let $\mathcal{KB} = (\mathcal{E}, \mathcal{R}, \mathcal{D})$ be a knowledge base. Let $\mathcal{E} \subset \mathcal{KB}$ be the set of available entities (ontological concepts). Let $\mathcal{I} \subset \mathcal{E}$ be the set of known incidents. Let $\mathcal{M} \subset \mathcal{E}$ be the set of known measures. Let $\mathcal{R} \subseteq \mathcal{E} \times \mathcal{E}$ be the relations between elements of $\mathcal{E}$. Let $\mathcal{D}$ be the set of available documents $\mathcal{D} = \{d_1, .., d_j\}$. Let $\mathcal{U} = \{(u \in d_i) | d_i \in \mathcal{D}\}$ be the set of available information units contained in D and $\mathcal{T}$ the set of terms used to textually build them.*

It is very important for the assessment of safety measures to pay respect to the *context* under that they are applied. For instance to supply rescue boats on a festival F_1 besides a river makes sense but for a festival F_2 in the forest it is totally senseless. The context are the factual parameters of the environment. If the parameters change, the relevant incidents, the according measures and the importances of both change. Respecting the documentary corpus $\mathcal{D}$, the context is for instance represented by certain parameters whose fulfillment is mentioned in the content of each document or the parameters, all documents have in common.

Definition 2. *Let $\mathcal{KB} = (\mathcal{E}, \mathcal{R}, \mathcal{D})$ be the knowledge base. For an entity $e_i \in \mathcal{E}$ let $C_{e_i} \subseteq \mathcal{E} \setminus \{e_i\}$ be the context of e_i with $C_{e_i} = \{c_1, .., c_j\}$.*

For instance, for the previous entities F_1 and F_2 the context C_{F_1} would be *near river* and C_{F_2} *in the forest*. We want to consider relations making entities a preventive or reactive measure to incidents. This means to focus on the chronological order of the execution. In some domains measures are classified into *before*, *during* and *after* an incident. We consider the relational classes *during* and *after* as unified. A measure that is taken before an expected incident is a preventive measure, a measure that is taken during or after an incident is a reactive measure. A measure may be of preventive as well of reactive character.

Definition 3. *Let $R_{PM-C} \subseteq M \times I$ be a relation under a context C, indicating which measures are taken in this context C before an incident, making them preventive measures. Let $R_{RM-C} \subseteq I \times M$ be the analogous relation, indicating which measures are taken after the occurring of an incident making them reactive measures.*

Additionally we rely on an *importance* ranking of incidents and measures under a given context. The importance is quantified by assigning a value between 1 for *important* and 0 for *not important*.

Definition 4. *Let $IMP(i, C) \in]0, 1]$ be the importance of an element of $\mathcal{I}$ under the context C. Let $IMP(m, i, C) \in]0, 1]$ be the importance of a measure m for the incident i under the context C.*

For a given context and relevant incident induced by the context the according measures are ordered by importance and classified into preventive and reactive measures. Altogether they build a kind of facilitated process snippet we call *PIRI (Preventive-Incident-Reactive-Interrelation)*. The presented model simplifies the real world for facilitation of assessment. Typically there is a cascade of measures that are executed in a specific order, e.g. in case of fire first evacuate all people, then close the doors and windows. A PIRI-snippet is formally defined as follows.

Definition 5. *Let $\mathcal{KB} = (\mathcal{E}, \mathcal{R}, \mathcal{D})$ be the knowledge base. Let $i \in \mathcal{I}$ be an incident and $C \subseteq \mathcal{E}$ a context. Then, we define a PIRI-snippet $PIRI(i, C) = \{C, PG\}$, where PG = (N,E) is a directed graph. We call PG the PIRI graph with nodes $N = C \cap \mathcal{M} \cup \{i\}$ all measures mentioned by C and the incident i and edges $E = \{R_{PM-C} \cup R_{RM-C}\} \cap \{i\}$ all edges containing the node i. The graph is weighted with node-weights IMP(N,C) and edge weights IMP(E,i,C).*

Figure 1 shows a PIRI-diagram for the incident I_1 under the context C consisting of several contextual elements $(c_1, c_2, c_3, ..)$.

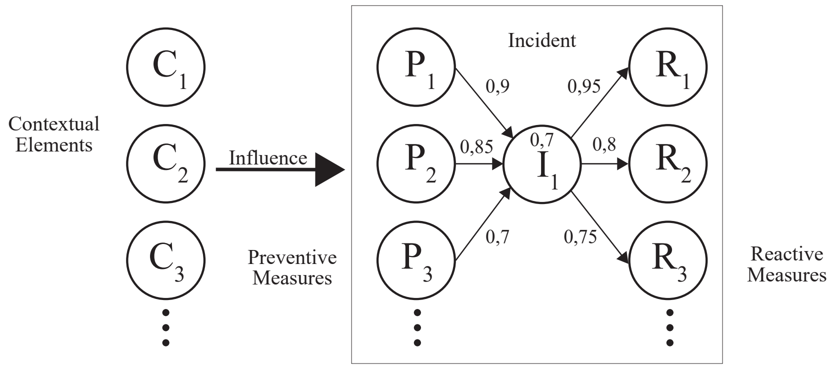

Fig. 1: PIRI-diagram under a context $C = (C_1, ..., C_j)$ showing the ranked preventive and reactive measures with the according importance weight.

2.1 Ontological Representation of the PIRI-Structure

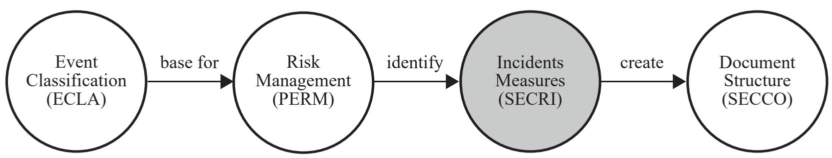

Fig. 2: Extension of the document creation workflow by the SECRI-ontology.

We extend a previously introduced ontology [9] by the definition of incidents, measures and the PIRI-snippet. The existent ontology was used for the classification of public events (O_{ECLA}) and the structuring of the according regulatory documents (O_{SECCO}) as depicted in Figure 2. For the ontological description of incidents we now continued the elaboration of the SECRI ontology (O_{SECRI}). The ontology describes the hierarchical context of incidents with a focus on public events. We use the SKOS ontology [17] and the PROV ontology [11] as upper ontologies. The SKOS ontology provides knowledge formalization and structuring capability. The PROV ontology supports the representation of provenance information to model the multi-agent-character of the scenario which is induced by the involvement of several authors and addressees. For the ontological implementation of a PIRI-snippet we introduce the analogous classes and interweave them with the documentary structure. An information unit is represented as a *secri:InformationUnit*. This passage of text semantically targets a *secri:Incident* or *secri:Measure* and is part of a document represented by *secco:Document*. Incidents and measures are subsuming classes as the top of a hierarchy. In the case study we will see an example of this hierarchy in the domain of public events. A graphical excerpt of the ontology can be seen in the Figure 3.

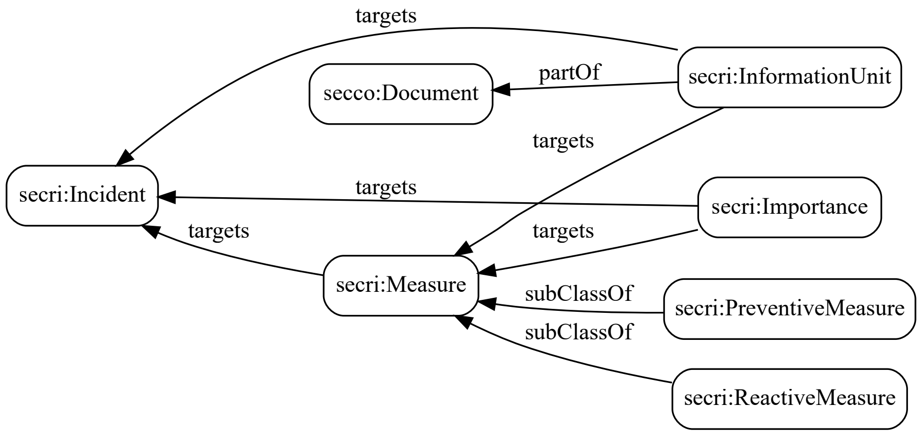

Fig. 3: Class representation of a PIRI-snippet.

2.2 Case-Based Representation of the PIRI-Structure

A convenient case-based representation for the so far described scenario internalizes the document description of incidents and measures. For each incident mentioned by the regulatory document the preventive and reactive measures are combined into a PIRI-snippet. We choose a structural case representation using attributes and their values [5].

Definition 6. *A case* $c_1 = (d_1, l_1)$ *is defined by the incident and its context as problem description* $d_1 = \{C_1, i_1\}$ *and its solution* $l_1 = \{(C_1), (m \in R_{PM-C_1} \cap i_1), (m \in R_{RM-C_1} \cap i_1)\}$*, the combination of measures targeting the incident* i_1 *under a context* C_1*, separated into preventive and reactive measures.*

The problem descriptions and the solutions are conjunctions of elements of the knowledge base. The context may be replaced for a unique identifier naming the context without citing every component. For instance $C_1 = RegDocument1$ then $d_1 = \{RD_1, RainStorm\}$ and $l_1 = \{(RD1), (WeightTents \wedge GetForecast), (CloseLiquidGas \wedge LockDoors \wedge GetRainCoat \wedge Evacuate)\}$.

Definition 7. *The case base* $CB = \{c_1, ..., c_m\}$ *is the collection of all cases* c_i *extracted from available regulatory documents and constructed as described before as PIRI-snippets. A query* q *to the case base is a conjunct subset of (negated) measures and incidents.*

For instance, the query $q_1 = CloseDoors \wedge \neg LockDoors \wedge Evacuation \wedge RainStorm$ retrieves all other PIRI-snippets containing an evacuation and a closing and not *locking of doors*.

To retrieve cases, we search the case base for similar problem descriptions d_i for the query q_1. To define a *similarity function*, we consider all preventive measures, reactive measures and the incident as individual sub-parts. Each of these parts is then compared by a local similarity measure. With an aggregation function a global similarity measure is composed by weighting with the parameters $(\omega_P, \omega_I, \omega_R)$ and summed up as follows:

$$Sim_{\mathrm{PIRI}}(c_k, c_l) = \big[\,\omega_P Sim_P(P_k, P_l) + \omega_I Sim_I(i_k, i_l) + \omega_R Sim_R(R_k, R_l)\,\big]/3 \quad (1)$$

The incidents and measures are classified by a taxonomy that was derived from the connected ontology, building the base for the similarity assessment and adaptation. The local similarities Sim_P, Sim_I, Sim_R are calculated via the taxonomic order of its elements. The incidents I and the measures P,R are hierarchically structured. Each element of the hierarchy is assigned with a likelihood symbolizing the similarity of its sub-elements. The similarity of the leaf-elements is set to *1* and to *0* for the root element. The similarity increases with depth d of the element according to for instance $sim_d = 1 - 1/2^d$ [1]. If we want to compare two PIRI-snippets it is desirable to consider the context. For this reason we define the following extended similarity measure under the context C:

$$Sim_{\mathrm{Context}}(c_k, c_l) = \big[\,\omega_1 Sim_{\mathrm{PIRI}}(c_k, c_l) + \omega_2 Sim_C(Context_k, Context_l)\,\big]/2 \quad (2)$$

The context may for instance be the fulfillment of a classification hierarchy describing the environmental parameters.

For instance, security measures under a context of high consumption of alcoholic beverages are to be considered different as under a context of low consumption of alcohol. So Sim_C is set to the similarity function used in that scenario weighted by the weights $\omega_i \in [0,1]$ working as biases.

2.3 Constrained-based Extension

The importance ranking can also be used as an order of execution of measures. The most important measures have to be taken first. But sometimes less important measures have to be taken before other, more important measures. This pays attention to the so called *concatenation of circumstances.* It is necessary to introduce a (partial) order of measures additionally to the order induced by *preventive* and *reactive* and the importance ranking.

Definition 8. *For two measures m_1 and m_2 the constraint $m_1 \prec m_2$ states that m_1 should be taken before m_2.*

An obvious problem is as follows. To avoid theft or unauthorized access especially large buildings have to be locked after an evacuation. This can yield people being locked inside the building. In reality it is often too complex or not possible to describe for each incident an order of taking the measures. Additionally in a multi-agent-scenario it is very difficult to execute instructions being too complex or too numerous. We therefore take a simple strategy of providing only rules for pairwise measures, as described before (*Evacuate* $\prec$ *LockDoors*).

3 Case Study

We exemplify the previous approach by a case study in the domain of public events. We started with 15 regulatory documents in the domain of public events that were annotated manually by three different domain experts. This corpus is extended as a basis for the present evaluation. For the annotation process we developed and evaluated several ontologies. These were used for the classification of public events (O_{ECLA}) and the structuring of the according security documents (O_{SECCO}). The following Table 1 shows the number of ontological concepts covered by each ontology.

Table 1. Number of ontological concepts for each ontology used.

O_{SECCO} Structuring	O_{ECLA} Classification	O_{SECRI} Incidents	O_{SECRI} Measures
278	136	115	72

3.1 Retrieval of Similar PIRI-Snippets

For the ontological description of security incidents we continue the elaboration of the SECRI ontology (O_{SECRI}). The SECRI ontology describes the hierarchical context of security incidents for public events. An excerpt of the ontology can be seen in Figure 4. In this work we extended the existing ontology by the capacity of modeling preventive and reactive measures for security incidents in the domain of public events. All ontologies where implemented using the semantic wiki KnowWE [4]. Amongst others we introduce the new classes *secri:Measure* as well as *secri:PreventiveMeasure* and *secri:ReactiveMeasure* as subclasses of *secri:Measure*.

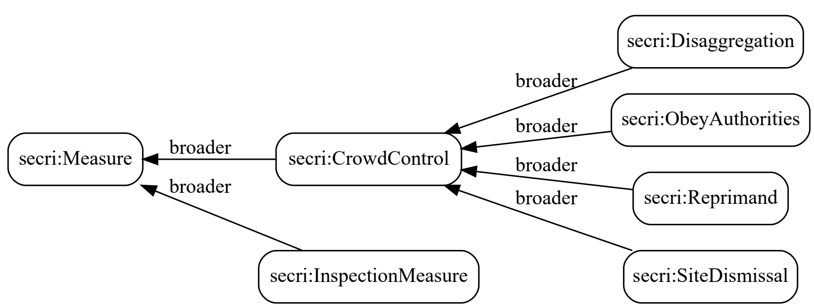

Fig. 4: Excerpt of the SECRI Ontology showing 7 of 72 measures.

For the case-based implementation we made use of myCBR [2]. The hierarchically structured incidents and measures represented in the SECRI ontology were exported to a myCBR model. The case-based attributes were arranged into taxonomies as local similarity measures. Those were aggregated into a global similarity measure for the assessment of the according PIRI-snippets. A number of relevant cases was extracted out of the corpus and installed in myCBR making up the experimental case base. Table 2 shows the number of different cases contained in the case base.

Table 2. Overview of different cases.

Different Contexts	PIRI-Cases	(Measures under Context)-Cases	Information Units
15	300	1500	500

To evaluate the similarity assessment induced by the PIRI-strategy we constructed a *post mortem analysis*. This means to take every case of the case base and use it as a query to the same case base. Our similarity measures are constructed symmetrically, consequently the query is commutative. As context we use the event classification ontology O_{ecla}. The context of each PIRI-snippet is represented by the factual parameters classifying the event extracted out of the according regulatory document. The pairwise similarities of the event classification cases are already available due to a post mortem analysis done in previous work [9]. Each document of the corpus mentions about 20 different incidents. We now focus on the incident *FireAndExplosion*. For this incident we pairwise calculate the similarity of the according PIRI-snippets. Afterwards we *apply* the context and calculate the context dependent similarity. Figure 5 shows for each pair of documents the similarity of the PIRI-snippet for the incident *FireAndExplosion* as well as the PIRI-similarity combined with the context. This comparison makes clear, where the influence of the context changes the similarity ranking of retrieved PIRI-cases.

3.2 Generation of Abstracted Information Units

In the following we present the strategy for the textual construction of PIRI-snippets and their adaptation for reuse. Figure 6 shows the workflow of breaking documents into reusable information units. Beginning with selected features it shows how they can be put together to form a new document. It presents which methods are used on each level for extraction, retrieval and adaptation. The relevant textual elements were extracted from the corpus and transferred into the ontological structures. The next step is to find the context dependent information and replace it to make them reusable. We therefore searched for elements of the domain vocabulary. Everything left we considered normal text or context related information that can be abstracted. A strategy for abstraction is to replace words by their class name. For instance a city name is replaced by *location data* or by the part-of-speech class. The following exemplary text for the incident *storm* shows, how an according passage of a security document would look in reality.

	christm PIRI0		wine PIRI1		wine PIRI2		folk PIRI3		city PIRI4		carne PIRI5		folk PIRI6		music PIRI7		carne PIRI8		fair PIRI9		fair PIRI10		running PIRI11		camp PIRI12		arena PIRI13		campus PIRI14	
PIRI0	x	x	0,5	0,6	0,6	0,6	0,5	0,5	0,6	0,6	0,5	0,5	0,3	0,5	0,3	0,5	0,4	0,5	0,3	0,5	0,3	0,5	0,3	0,3	0,3	0,3	0,4	0,5	0,0	0,2
PIRI1	0,5	0,7	x	x	0,9	0,8	0,4	0,5	0,3	0,4	0,4	0,5	0,6	0,7	0,3	0,3	0,6	0,5	0,4	0,5	0,4	0,5	0,5	0,4	0,3	0,3	0,4	0,4	0,3	0,3
PIRI2	0,5	0,5	0,9	0,8	x	x	0,5	0,5	0,4	0,4	0,5	0,4	0,6	0,6	0,3	0,3	0,5	0,5	0,4	0,6	0,4	0,5	0,4	0,4	0,3	0,4	0,4	0,4	0,2	0,3
PIRI3	0,5	0,5	0,4	0,5	0,5	0,5	x	x	0,5	0,6	1,0	0,8	0,4	0,5	0,3	0,5	**0,2**	**0,5**	0,4	0,5	0,4	0,4	0,2	0,3	0,2	0,3	0,6	0,6	0,1	0,2
PIRI4	0,6	0,6	0,3	0,4	0,4	0,4	0,5	0,6	x	x	0,5	0,6	0,3	0,4	**0,2**	**0,5**	**0,1**	**0,4**	0,1	0,3	0,1	0,2	0,2	0,3	0,1	0,2	0,7	0,6	0,0	0,1
PIRI5	0,5	0,5	0,4	0,5	0,5	0,4	1,0	0,8	0,5	0,6	x	x	0,4	0,5	0,3	0,5	**0,2**	**0,5**	0,4	0,5	0,4	0,5	0,2	0,3	0,2	0,2	0,6	0,5	0,1	0,2
PIRI6	0,3	0,5	0,6	0,7	0,6	0,6	0,4	0,5	0,3	0,4	0,4	0,5	x	x	0,3	0,4	0,3	0,5	0,5	0,5	0,3	0,4	0,3	0,3	0,3	0,3	0,4	0,5	0,1	0,3
PIRI7	0,3	0,5	0,3	0,3	0,3	0,3	0,3	0,5	**0,2**	**0,5**	0,3	0,5	0,3	0,4	x	x	0,3	0,5	0,3	0,2	0,3	0,3	0,1	0,3	0,3	0,3	0,3	0,4	0,0	0,1
PIRI8	0,4	0,5	0,7	0,6	0,6	0,5	**0,2**	**0,5**	0,8	0,8	**0,2**	**0,5**	0,3	0,5	0,3	0,5	x	x	0,5	0,5	0,5	0,5	0,6	0,5	0,4	0,4	0,2	0,3	0,1	0,1
PIRI9	0,3	0,5	0,4	0,5	0,4	0,6	0,4	0,5	0,1	0,3	0,4	0,5	0,5	0,6	0,3	0,2	0,5	0,5	x	x	0,9	0,8	0,2	0,3	0,3	0,3	0,2	0,3	0,2	0,3
PIRI10	0,3	0,5	0,4	0,5	0,4	0,5	0,5	0,5	0,1	0,2	0,5	0,5	0,3	0,4	0,3	0,3	0,5	0,5	0,9	0,8	x	x	0,2	0,3	0,3	0,3	0,2	0,3	0,2	0,3
PIRI11	0,3	0,3	0,5	0,4	0,5	0,4	0,2	0,3	0,2	0,3	0,2	0,3	0,3	0,3	0,1	0,3	0,6	0,5	0,2	0,3	0,2	0,3	x	x	0,3	0,3	0,3	0,4	0,3	0,3
PIRI12	0,3	0,3	0,3	0,3	0,3	0,4	0,2	0,3	0,1	0,2	0,2	0,2	0,3	0,3	0,3	0,3	0,4	0,4	0,3	0,3	0,3	0,3	0,3	0,3	x	x	0,1	0,2	**0,0**	**0,3**
PIRI13	0,4	0,5	0,4	0,4	0,4	0,4	0,6	0,6	0,7	0,6	0,6	0,5	0,4	0,5	0,3	0,4	0,2	0,3	0,2	0,3	0,2	0,3	0,3	0,4	0,1	0,2	x	x	0,2	0,3
PIRI14	0,0	0,2	0,3	0,3	0,2	0,3	0,1	0,2	0,0	0,1	0,1	0,2	0,1	0,3	0,0	0,1	0,3	0,3	0,2	0,3	0,2	0,3	0,3	0,3	**0,0**	**0,3**	0,2	0,3	x	x

Fig. 5: Post mortem analysis of the PIRI-snippets for the measure *FireAndExplosion* without and with the context of the event classification. The values show the similarities $\mathrm{Sim}_{\mathrm{PIRI}} \mid \mathrm{Sim}_{\mathrm{Context}}$. The value for $\mathrm{Sim}_{\mathrm{Context}}$ was calculated out of $\mathrm{Sim}_{\mathrm{PIRI}}$ and $\mathrm{Sim}_{\mathrm{ECLA}}$ which were weighted with 0.5 each. A significant change of the retrieval by the incorporated context is marked bold.

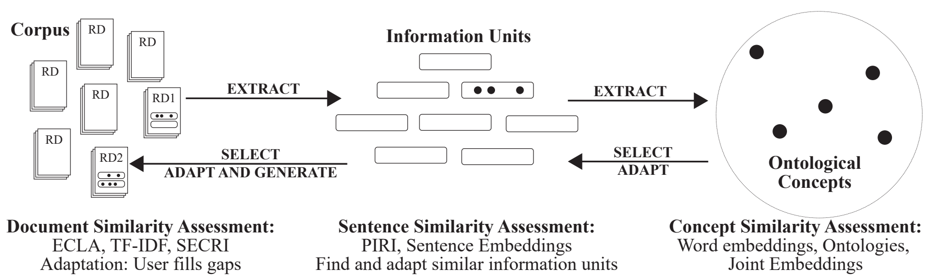

Fig. 6: Workflow of decomposing and recomposing security documents.

"Storm. Get weather information on a regularly basis from the munich weather station. Weight all tents with heavy material or fix with ropes. In case of upcoming storm, evacuate the event site using the franz josef avenue and call the fire department."

The PIRI-snippet with exemplary importance values for this would be:

Preventive(*WeightTents(0.9),GetWeatherForecast(0.8)*)

Incident(*Storm*)

Reactive(*CallFireDepartement(0.9),FullEvacuation(0.8)*).

An abstracted information unit for the measure *FullEvacuation* would be:

"[FullEvacuation][StopWord][EventSite][Verb][StopWord][LocationData]"

This information unit can be adapted for instance to the measure *PartialEvacuation*. The ontological concept *FullEvacuation* is replaced by a retrieved information unit for the new measure. The concept *EventSite* is for instance replaced by the more specific concept *EventSiteComponent*. This information can be retrieved out of other cases because *PartialEvacuation* is commonly combined with *EventSiteComponent*. The concept *LocationData* has to be replaced by the contextual location information which is left to the user. The stop words are inserted and corrected by a natural language generation tool or the user. The generated textual passage before stop word correction and context correction looks as follows:

"[Partial evacuation of the affected area][the] [EventSiteComponent][using][the][LocationData]"

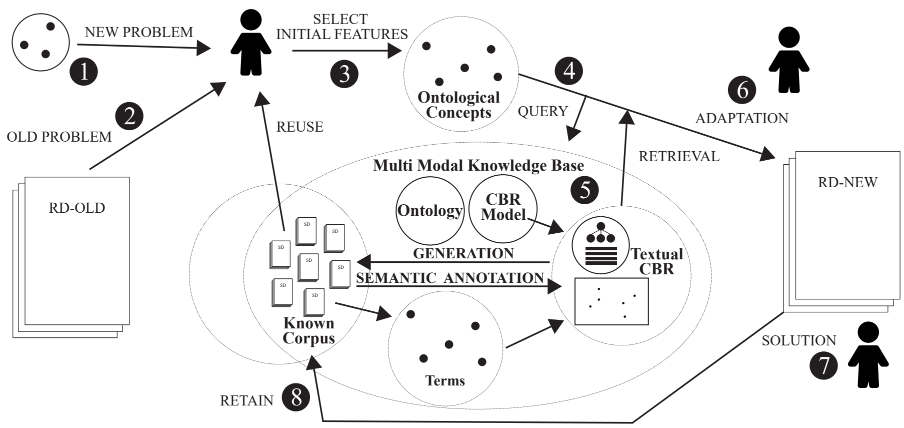

Fig. 7: Case-based cycle of regulatory document assessment.

Figure 7 shows the user interaction and the case-based cycle of natural document extraction and generation. In step (1) a new problem arises. That may be for instance that a new regulatory document is required or an existing document has to be improved as shown in step (2). All features are extracted out of the problem description and the old document at step (3) and queried to the knowledge base at step (4). The retrieved features, phrases and documents are returned in step (5) and adapted in step (6) which requires user interaction. The new regulatory document is used (7) and retained in the corpus enlarging the case base (8).

3.3 Results and Discussion

The results of the case study for the retrieval of similar information units are very promising even without user support. The incorporation of the contextual paradigm significantly improved the simulation of the real world scenario. Regarding the generation of regulatory documents the results were quite good when supported by the user. To answer the *initial question*, which incidents are likely to happen, the context-based assessment can be used - similar context points to similar incidents. Same holds for the measures that are suitable for an incident. The importance of incidents and measures is made accessible by the percentage of cases covering the incident or measure under a certain context.

4 Related Work

We started the research for related work to this paper with an overview of state of the art publications in the domain of *natural language generation* presented by Gatt and Krahmer [7]. Most of the presented work requires a large corpus for the application of statistical methods. More suitable for our necessities seemed grammar-based approaches. This lead us to the idea of abstracting text by giving it a pseudo grammar structure.

There exists some work for the assessment of incidents in different domains. A similar approach we want to mention was presented by Sizov et al. [14]. The work focuses on the extraction and the (case-based) adaptation of explanations contained in incident reports in the transportation domain. The work differs in that way that we are aiming for a holistic document oriented and ontology-based approach with user support for generation. A framework for the connection of ontologies and constraints for the assessment of workflows was presented by Nguyen et al. [12].

The structural integration of context into the case-based assessment was covered by various authors. Different approaches for the incorporation of context into a case-based decision were proposed by Pla et al. [13]. We adapted the method of context stacking for this scenario. A conceptual revision of the context-based reasoning paradigm was presented by Stensrud et al. [15]. For the role of context in case-based reasoning a good work was published by Khan et al. [8] as well as by Craw and Aamodt [6] for the use of similar case clusters for representing context. The ontological side of context representation was for instance presented in a thoughtful way by Strang et al. [16] and Xu et al. [18].

5 Conclusions

In this paper we presented a data structure called PIRI for the representation of a regulatory document describing incidents and according measures. After formally describing it, we transferred the structure into a case-based model. Using this model an approach was shown for the adaptation of similarity measures to different context. In a case study the approach was applied to a corpus of regulatory documents of the domain of public events. What we left for future work are strategies for the identification of relevant attributes out of existing cases. The application of attribute dependent weights would help to individually adjust the influence of the context onto the case-based assessment. Additionally, in the field of document generation the integration of grammar-based natural language generation approaches seems to be promising. To adapt abstracted information units to different contexts would help to reduce the needed user support.

References

1. Bach, K., Althoff, K.D.: Developing case-based reasoning applications using myCBR3. In: Agudo, B.D., Watson, I. (eds.) Case-Based Reasoning Research and Development. pp. 17–31. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
2. Bach, K., Sauer, C., Althoff, K.D., Roth-Berghofer, T.: Knowledge modeling with the open source tool myCBR. In: Proceedings of the 10th International Conference on Knowledge Engineering and Software Engineering - Volume 1289. pp. 84–94. KESE'14, CEUR-WS.org, Aachen, Germany (2014)
3. Baumeister, J., Reutelshoefer, J.: The connectivity of multi-modal knowledge bases. CEUR Workshop Proceedings **1226**, 287–298 (01 2014)
4. Baumeister, J., Reutelshoefer, J., Puppe, F.: KnowWE: A semantic wiki for knowledge engineering. Applied Intelligence **35**(3), 323–344 (2011)
5. Bergmann, R.: Experience Management. Springer, Berlin, Heidelberg (2002)
6. Craw, S., Aamodt, A.: Case-Based Reasoning as a Model for Cognitive Artificial Intelligence: 26th International Conference, ICCBR 2018, Stockholm, Sweden, July 9-12, 2018, Proceedings, pp. 62–77 (07 2018)
7. Gatt, A., Krahmer, E.: Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. CoRR **abs/1703.09902** (2017)
8. Khan, N., Alegre, U., Kramer, D., Augusto, J.C.: Is 'context-aware reasoning = case-based reasoning'? In: Brézillon, P., Turner, R., Penco, C. (eds.) Modeling and Using Context. pp. 418–431. Springer International Publishing, Cham (2017)
9. Korger, A., Baumeister, J.: The SECCO ontology for the retrieval and generation of security concepts. In: Cox, M.T., Funk, P., Begum, S. (eds.) ICCBR. Lecture Notes in Computer Science, vol. 11156, pp. 186–201. Springer (2018)
10. Korger, A., Baumeister, J.: Textual case-based adaptation using semantic relatedness - a case study in the domain of security documents. In: Wissensmanagement Potsdam (2019)
11. Moreau, L., Groth, P.: Provenance: An Introduction to PROV. Synthesis Lectures on the Semantic Web: Theory and Technology, Morgan and Claypool (2013)
12. Nguyen, T.H.H., Le-Thanh, N. (eds.): Ensuring the Semantic Correctness of Workflow Processes: An Ontological Approach. KESE 2014 Knowledge Engineering and Software Engineering (2014)
13. Pla, A., Coll, J., Mordvaniuk, N., López, B.: Context-aware case-based reasoning. In: Prasath, R., O'Reilly, P., Kathirvalavakumar, T. (eds.) Mining Intelligence and Knowledge Exploration. pp. 229–238. Springer International Publishing, Cham (2014)
14. Sizov, G., Ozturk, P., Marsi, E.: Let me explain: Adaptation of explanations extracted from incident reports. AI Communications **30**, 1–14 (06 2017)
15. Stensrud, B.S., Barrett, G.C., Trinh, V.C., Gonzalez, A.J.: Context-based reasoning: A revised specification. In: FLAIRS Conference (2004)
16. Strang, T., Linnhoff-Popien, C., Frank, K.: Cool: A context ontology language to enable contextual interoperability. vol. 2893, pp. 236–247 (01 2004)
17. W3C: SKOS Simple Knowledge Organization System Reference: http://www.w3.org/TR/skos-reference (August 2009)
18. Xu, N., Zhang, W., Yang, H., Zhang, X., Xing, X.: Cacont: A ontology-based model for context modeling and reasoning. Applied Mechanics and Materials **347-350** (03 2013)