Implementing Semantic Correlation Rules in the Domain of Energy Technology for Transportation Systems

Anna-Maria Kruse[1,*], Celine Wiedewilt[1,**] and Lars Tempelfeld[1,***]

Author Affiliations
[1] Karlsruhe University of Applied Sciences, Faculty of Information Management & Media
Amalienstraße 81-87
76133 Karlsruhe
Germany

Author Emails
* *kran1035@h-ka.de*
** *wice1011 @h-ka.de*
*** *tela1014@h-ka.de*

Abstract. This research explores the correlation between topics in a content delivery portal (CDP) and spare parts catalogue (SPC) instances, focusing the domain of energy technology for transportation systems. It is based on use cases that require further textual content in the CDP as well as reference links to corresponding instances in the SPC. Furthermore, this paper aims to show the process of creating Semantic Correlation Rules (SCR) – rule-based links between instances in a semantic network – based on the previously elaborated ontology. Moreover, it aims to assess, whether SCR meet users' needs and discusses the strengths and difficulties of this implementation approach. In summary, this paper shows the general implementation process of SCR as well as the associated requirements in the field of energy technology for transportation systems.

1 INTRODUCTION

The importance of semantic modelling technologies in the field of information management and technical communication has increased significantly in the recent years. Elaborated content management is still essential, but the focus has now shifted to providing contextual results tailored to users' needs and the challenges associated with this. Addressing these challenges requires the application of intelligent metadata. Concise metadata and their consistent use are essential to meet users' demands. In addition, the recent demand has led to dynamic information delivery with its implementation in CDPs.

Data originating from the field of technical communication is often maintained in different data stores. To solve the challenge of insufficient or excessive information – and to ensure users of technical information receive concise information within a specific time frame and in an appropriate context – the different stored data must be linked dynamically. One approach for this challenge is the creation of microDocs that are realized by so called Semantic Correlation Rules. These microDocs aim to deliver concise information within a logically given context by creating a smaller dynamic document with independent topics [1,2]. The dynamic documents can be implemented using SCR, which are based on relevant use cases that specify the correlations between instances, e.g. independent topics. User-driven use cases of information needs and implicit knowledge requests build the foundation of SCR due to their transformation of implicit to explicit knowledge. The development process includes the merge of an ontology with instances originating from these various data stores. A rule-based correlation between instances in this ontology can then be used to generate dynamic links in the respective CDP that offer further information in the form of topics.

2 SUBJECT OF RESEARCH

Recently, microDocs and SCR have been introduced in the field of technical communication. SCR are a concept in which metadata originating from various data domains is linked via sets of logical rules (InRules and OutRules). These rules describe the start and end points of information request and logical information delivery. The relations enable the automated aggregation of microDocs, which can be seen as a set of topics. The utilization of SCR includes standard semantic formats; therefore, no specific authoring system is required [2].

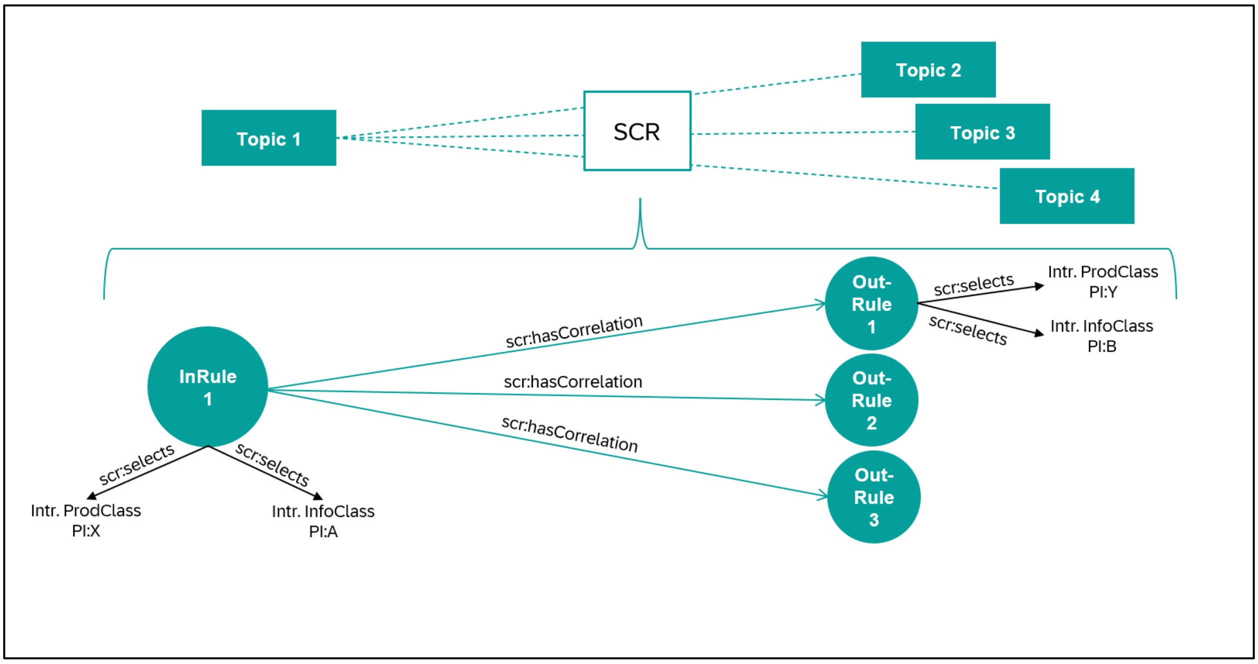

FIGURE 1. The concept of SCR [2]

In figure 1 the concept of SCR is displayed: Topic 1 is correlated to Topic 2,3 and 4 via SCR, while use cases describe which topics are linked to each other. SCR consist of a set of In- and OutRules. Figure 1 also shows the basic situation of the SCR implementation concept – the InRule selects two intrinsic metadata: the specific product (f.e. X) and information class (f.e. A) according to PI-Class. Due to its correlation to the OutRule, which also selects intrinsic product and information class, links between topics using these metadata values can be generated dynamically. This applies to OutRule 2 and 3 as well.

3 METHOD

For this research, use cases and data were provided by a cooperating partner, based in the field of energy technology for transportation systems. It focusses on topics where metadata has been applied according to the PI-Class method in parts in combination with the iiRDS – an "intelligent information request and delivery standard". The associated metadata from the technical documentation represents the high product variance and the encountered information classes to distinguish between various information types. To link metadata that originates from different data stores, a domain ontology must be created. Details of the development process are described in [3]. The merged ontology can be accessed via a graph database, which is usually created using labelled property graph or Resource Description Framework (RDF) and Ontology Web Language (OWL). SPARQL can then be used to query certain information from the ontology. Information exchange between applications is facilitated by a REST API server with system-specific requests and SPARQL queries [4]. The rules can be tested using SPARQL queries and SWAGGER and lastly in the productive content provisioning systems CDP and SPC. A detailed description of the technical implementation can be found in [5].

Building the Ontology

In order to implement the correlation between metadata originating from different data stores, an ontology has to be created, where in this research project metadata from the Component-Content-Management-System (CCMS) and metadata from the SPC are merged in a domain ontology. It consists of main classes (content-engineering and product-engineering) and their subclasses. The class "content-engineering" uses a class-based approach, whereas the "prod-engineering" class uses an instance-based approach for metadata storage. A detailed description of the advantages of each approach can be found in [3]. The modelling was done in the open-source application Protégé.

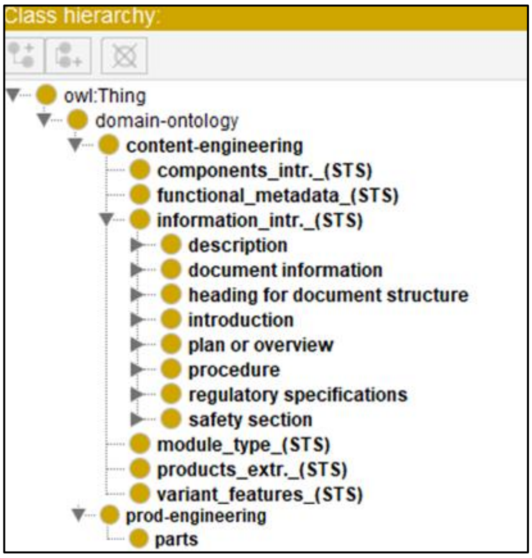

FIGURE 2. Excerpt of the domain ontology developed in Protégé.

In this domain ontology, metadata originating from different data stores as explained in the previous, can be linked via SCR.

Collecting Use Cases

To implement SCR, a number of use cases have been collected in a standardized excel sheet [6] as shown in figure 3. The rows contain the use cases, and the columns the specific attributes of each use case, for example: the use case name, the system type start and end, and the metadata that defines the specific use case. Within this project four different types of use cases according to their system type start and end (CDP to CDP, CDP to SPC, SPC to SPC, SPC to CDP) could be identified. For example, if the user starts with the topic "transport without packaging" (metadata from intrinsic information class according to PI-Class) in the system CDP, a link to the topic "general safety" should be provided to the user. The intrinsic component class remains unchanged – "total product".

 https://doi.org/10.1051/shsconf/202419402005

General information			Starting point			Target point	
Use case ID	Use case name	Use case description	System type start	List of metadata start	InRule name	System type target	List of metadata target
1	Transport without packing		CDP	components_intr. (STS): total product; information_intr. (STS): transport without packing; products_extr. (STS): SBE; WBL-Z;	InRule0001_Transport_without_packing	CDP	components_intr. (STS): total product; information_intr. (STS): general safety; products_extr. (STS): SBE; WBL-Z;
1	Transport without packing		CDP	components_intr. (STS): total product; information_intr. (STS): transport without packing; products_extr. (STS): SBE; WBL-Z;	InRule0001_Transport_without_packing	CDP	components_intr. (STS): total product; information_intr. (STS): transport safety; products_extr. (STS): SBE; WBL-Z;
1	Transport without packing		CDP	components_intr. (STS): total product; information_intr. (STS): transport without packing; products_extr. (STS): SBE; WBL-Z;	InRule0001_Transport_without_packing	CDP	components_intr. (STS): total product; information_intr. (STS): packing; products_extr. (STS): SBE; WBL-Z;
1	Transport without packing		CDP	components_intr. (STS): total product; information_intr. (STS): transport without packing; products_extr. (STS): SBE; WBL-Z;	InRule0001_Transport_without_packing	CDP	components_intr. (STS): total product; information_intr. (STS): storage conditions; products_extr. (STS): SBE; WBL-Z;
1	Transport without packing		CDP	components_intr. (STS): total product; information_intr. (STS): transport without packing; products_extr. (STS): SBE; WBL-Z;	InRule0001_Transport_without_packing	CDP	components_intr. (STS): total product; information_intr. (STS): attachment; products_extr. (STS): SBE; WBL-Z;
2	Maintenance interval 2		CDP	components_intr. (STS): total product; information_intr. (STS): maintenance plan; products_extr. (STS): SBE; WBL-Z; functional_metadata (STS): maintenance interval 2;	InRule0002_Maintenance_interval_2	CDP	components_intr. (STS): total product; information_intr. (STS): general safety; products_extr. (STS): SBE; WBL-Z;
2	Maintenance interval 2		CDP	components_intr. (STS): total product; information_intr. (STS): maintenance plan; products_extr. (STS): SBE; WBL-Z; functional_metadata (STS): maintenance interval 2;	InRule0002_Maintenance_interval_2	CDP	components_intr. (STS): total product; information_intr. (STS): maintenance safety; products_extr. (STS): SBE; WBL-Z;
2	Maintenance interval 2		CDP	components_intr. (STS): total product; information_intr. (STS): maintenance plan; products_extr. (STS): SBE; WBL-Z; functional_metadata (STS): maintenance interval 2;	InRule0002_Maintenance_interval_2	CDP	components_intr. (STS): total product; information_intr. (STS): maintenance plan; products_extr. (STS): SBE; WBL-Z; functional_metadata (STS): maintenance interval 1;

FIGURE 3. Standardized excel sheet for the collection of use cases.

Visualizing the Ontology and Forming Queries with SPARQL

After modelling the domain ontology and collecting use cases, the use cases can be implemented in a graph database called "GraphDB" using SPARQL. SPARQL is a query language developed in the context of the semantic web and is used in modelling SCR. It enables the retrieval and customization of triple data stores. For the implementation in this research project the standard SCR modelling pattern is used. First, the InRules are defined. This requires a unique name in the different languages the company provides, as well as the metadata IDs of the starting point and the correlated OutRules. The following code snippet (figure 4) shows how the SCRs were inserted using SPARQL queries. Each InRule has a specific name – the name follows a specific pattern – and a type definition. The 'scr:selects' commands are required to select the correct metadata that can be identified by its unique ID number originating from the CCMS or SPC. The last paragraphs correlate the associated OutRule(s).

```
scr:InRule0001_Transport_without_packing rdfs:label "InRule0001 Transport without packing"@en.
scr:InRule0001_Transport_without_packing rdfs:label "InRule0001 Transportieren ohne Verpacken"@de.

scr:InRule0001_Transport_without_packing rdf:type scr:InRule.

scr:InRule0001_Transport_without_packing scr:selects sts:71261835.
scr:InRule0001_Transport_without_packing scr:selects sts:70144267.

scr:InRule0001_Transport_without_packing scr:hasCorrelation scr:OutRule0001_Transport_without_packing.
scr:InRule0001_Transport_without_packing scr:hasCorrelation scr:OutRule0002_Transport_without_packing.
scr:InRule0001_Transport_without_packing scr:hasCorrelation scr:OutRule0003_Transport_without_packing.
scr:InRule0001_Transport_without_packing scr:hasCorrelation scr:OutRule0004_Transport_without_packing.
scr:InRule0001_Transport_without_packing scr:hasCorrelation scr:OutRule0005_Transport_without_packing.
```

FIGURE 4. Definition of an InRule with SPARQL in GraphDB

In addition, the insert of the OutRules into the graph database works likewise – without the correlation. The OutRules select the metadata IDs of the targeted endpoint. After inserting the SCR, the technical implementation follows. The technical implementation of SCR contains generic SPARQL queries with variables in the REST API. For each request, the variables are filled dynamically with metadata from the topic currently viewed. The graph database then identifies the corresponding InRule and answers the request.

iiRDS Mapping Extensions

In this research project, the CDP from the cooperating company imports and exports metadata via iiRDS, not the initial CCMS metadata. This standard works with differing information classes, therefore, an intermediate step had to be added to the SCR concept to reverse the transformation within the graph database. For this, the SCR specification uses the DynaRule extensions invoking server-side functions in the REST API of the corresponding web-service. For example, in a first intermediate step, the InRule calls the DynaRule IfInfoMapping.PI2iiRDS (iiRDS:InfSubject) checking if the information subject of a CDP event transmitted via the http-request corresponds to the CCMS metadata in the ontology.

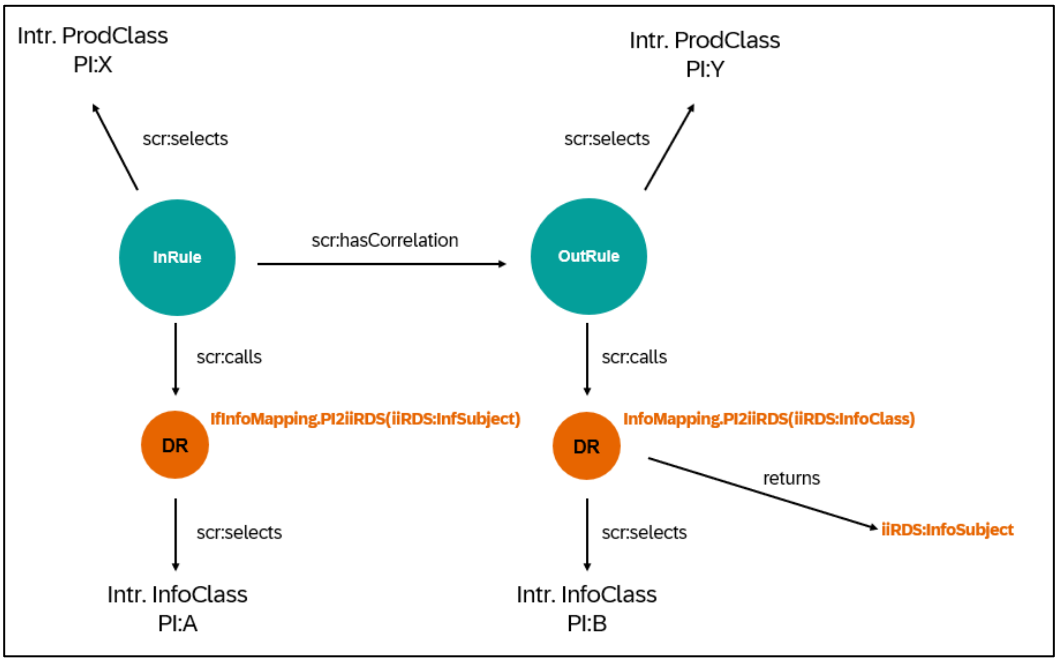

FIGURE 5. Extension of the SCR concept with DynaRules (DR) [7]

In figure 5, the orange dots illustrate the enhancement with additional functions that had to be implemented on the server-side. The function InfoMapping.PI2iiRDS (iiRDS:InfoClass) called by the OutRule ensures that the API returns an information subject covered by the iiRDS standard, thus information architects still can use the initial CCMS metadata to describe the use cases. Therefore, in this implementation, the DynaRules and functions are aware of the mapping mechanism between PI-Class and iiRDS [7].

4 REPRESENTATION IN SPC AND CDP

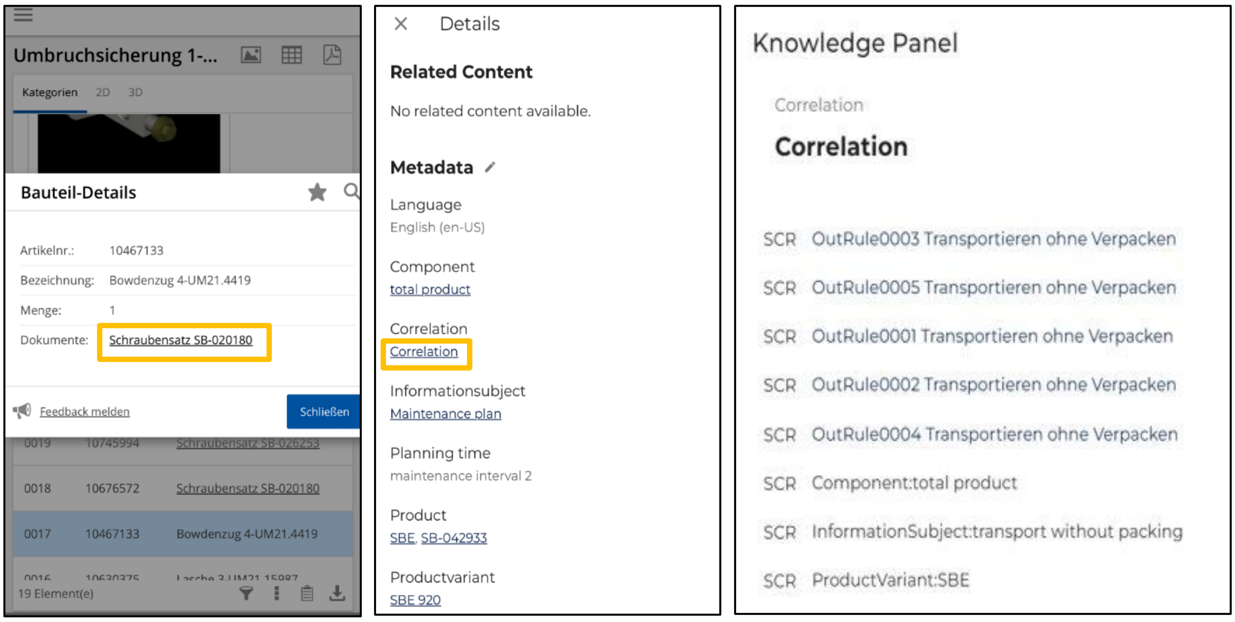

FIGURE 6a, b, c. In the SPC of Quanos Content Solutions, the correlations are shown in the form of deep links. In figure 6a this deep link correlates the spare part "Bowdenzug" (engl. tow rope) with the spare part "Schraubensatz" (engl. screw set). In addition, in the CDP from Empolis the user receives, if the highlighted link "Correlation" in Figure 6b is being used, correlated content in form of deep links, shown in figure 6c.

5 KEY-FINDINGS

In general, the implementation process of the SCR was achieved during the specified research period. However, some prior knowledge of databases and programming is required. The SCR presented in the form of deep links can be generated dynamically in the respective system to meet the requirements of the user.

- The standardized approach of using an excel template to collect use cases prevents the implementation of SCR from terminological inconsistencies and further challenges. Furthermore, it is possible to implement the use cases more effectively when using the template.
- SCR work for four different types of use cases as shown in the previous. A correlation of topics from two different systems (CCMS & SPC) is technically possible, but at this point of the research the quality of the resulting correlated topics and upcoming challenges were not the focus.
- Moreover, the SCR concept is not language dependent when using IDs instead of labels for the 'scr:selects' command. In general, the SCR concept is not bound to labels, but in the context of this study, the functionality of SCR is currently dependent on the labels of In- and OutRules.
- A prerequisite for the implementation of SCR is the standardized use of metadata, such as the PI-Class. In such a case, a standard approach for SPARQL and API development is straightforward. Additional metadata transformations to other approaches (e.g. iiRDS) can be covered by enhanced SCR modelling (1.2) as shown in the research by the DynaRules.

6 CONCLUSION AND OUTLOOK

With the SCR approach, the requirements of the users in technical communication can be met by providing further contextual information given by logical rules for metadata. The concept is not limited to the two project-specific systems, the implementation of SCR can be extended to any API-calling system and can provide a specification of its deep links. Moreover, the iiRDS mapping extension offers an advantage: information architects can focus solely on their domain knowledge without needing to understand mapping details.

However, some prospects can be considered for further development: In order to organize the collection of use cases more effectively, an elaborated standardized process could be evolved. Additionally, a further extension of this research and the specific industry data could be the extension of a concept that overcomes the use of language-dependent labels for In- and OutRules, for example, to avoid terminological inconsistencies. Another point to consider is the optimization of the link texts displayed in figure 7a, b, c regarding their user experience.

ACKNOWLEDGMENTS

This work was supported by Schunk Transit Systems and the information they provided. The authors gratefully acknowledge the helpful discussions with Marcus Bartel, Stephan Bartl, Sebastian Ries and Nino Tiezel (all Schunk Transit Systems). Furthermore, we acknowledge the technical support by Florian Rommel, Maurice Daum and Lena Wenner (all ICMS). We would also like to thank the system providers involved, namely Empolis and Quanos Content Solutions for the provision and adaption of their systems. Additionally, we want to thank the International Office of Karlsruhe University of Applied Sciences and the study program of Communication and Media Management for the financial support.

REFERENCES

1. W. Ziegler, *Extending intelligent content delivery in technical communication by semantics: microdocuments and content services.* Proceedings of the ETLTC ACM Chapter International Conference. Aizuwakamatsu, Japan. (2020) https://doi.org/10.1051/shsconf/20207703009
2. W. Ziegler, *Semantic Correlation Rules as a Logic Layer between Content Management and Content Delivery*. Proceedings of the ETLTC ACM Chapter International Conference. Aizuwakamatsu, Japan. (2021) https://doi.org/10.1051/shsconf/202110202007
3. L. Tempelfeld, C. Wiedewilt, A. Kruse, *Transforming Domain Data into Semantic Models for Content Delivery Applications.* Proceedings of the ETLTC ACM Chapter International Conference. Aizuwakamatsu, Japan. (2024)
4. F. Rommel and W. Ziegler, *A Microservice-based Infrastructure for the Academic Research on Semantic Knowledge Graphs and Derived Applications*, Proceedings of the ETLTC ACM Chapter International Conference. Aizuwakamatsu, Japan. (2022) https://doi.org/10.1051/shsconf/202213902002
5. M. Daum, J. Telatinski, P. Rannacher, W. Ziegler, *Technical Implementation of Semantic Correlation Rules via API programming for Basic and Extended Processing Logics.* Proceedings of the ETLTC ACM Chapter International Conference. Aizuwakamatsu, Japan. (2023) https://doi.org/10.1063/5.0183352
6. J. Telatinski, P. Rannacher, M. Daum, *Implementation of Semantic Correlation Rules for Multiple Use Cases in an Industrial Context.* Proceedings of the ETLTC ACM Chapter International Conference. Aizuwakamatsu, Japan. (2023) https://doi.org/10.1063/5.0184363
7. W. Ziegler, course work communication SIM class Karlsruhe University of Applied Sciences 2024