ABSTRACT
The search for existing solutions in mechanical engineering is a key-factor for successful product development. By reusing existing solutions, the amount of work and costs within the product development process can be reduced. However, not only exact matches for a required solution are helpful, but also the suggestion of similar solutions could often meet the requirements and bring benefit. Therefore, the presented research aims at improving the process of retrieving existing solutions similar to a user query. This is achieved by developing an abstraction level in form of a shape classification for the objects handled in a technical solution. An ontology is used for modeling the required concepts and providing the necessary relations to shape classification. Semantic similarity measures are applied for calculating the similarity of technical solutions and the user query according to the information modeled in the ontology. With the help of a developed prototype that uses the ontology as knowledge repository for annotating and searching solution documents, the presented approach is applied and evaluated in the field of automation industry.

Keywords: solution search, ontology development, similarity of technical solutions, object abstraction, object shape

1 INTRODUCTION
In product development, the reuse of existing knowledge about technical solutions may reduce the amount of development cycles and conception rework and therewith decrease the efforts of time and costs. Principally, several sources exist for supporting this knowledge. In a study in the German automation industry, sources for the search of existing solution knowledge were identified [1]. Besides direct personal communication, organization-internal knowledge sources (e.g. project folders or databases), construction catalogues, internet portals, and publically available marketing documents were identified as mostly used.

However, a person (e.g. an engineer) who wants to retrieve existing solution knowledge faces several general barriers (see Figure 1). First of all, solution knowledge is mostly unstructured in text files like emails or project reports and the access to unstructured data is often insufficient [2]. Secondly, different wordings are used by the involved people [3]. This hinders the access via a normal full-text search [4]. Furthermore, varying taxonomies and classifications of the involved divisions like sales, marketing, and engineering contribute to the barrier that hinders the access to needed solutions [5].

![Figure 1: General and specific barriers in the process of retrieving technical solutions](image-url)
In addition to these general barriers, a further problem hinders the access to needful knowledge about existing solutions. Often, similar solutions could also fulfill the engineer’s needs, but are not found as search result. For example, an engineer searching for a solution for transferring a bottle for packaging could also be satisfied by a solution for transferring a can. Probably, the engineer would have to adjust some boundary conditions and technical specifications of this existing similar solution. Nevertheless, the similar solution approach could be helpful and the amount of engineering work would certainly be less than developing a completely new solution for packaging a bottle.

Due to these existing problems, research is needed to develop methods and technologies for solution retrieval in the field of mechanical engineering. In the Use-Case PROCESSUS, which is one part of the German research project THESEUS, an ontology based solution search will be developed and evaluated that shall enable the quick and precise access to the required solution and application knowledge ([6], [7]). Therewith, the development processes will be fundamentally accelerated and optimized. PROCESSUS is specified and demonstrated in a scenario taken from the automation and packaging industry. In this scenario, an engineer in the design process should be supported in searching for solutions and storing knowledge about formerly developed and existing solutions.

For supporting this handling of solution knowledge, an ontology has been developed that is used for capturing the knowledge of technical solutions (for details about the development of the ontology and related approaches see [8] and [9]). In the ontology, the solution is described mainly by the performed functions. According to common design research theory, the functions are composed of an object and an operation performed on the object [10, 11]. Abstraction levels for operations are used for providing an abstract access to the technical functions [9]. With the help of a developed prototype, the ontology is used for the automated annotation of existing solution documents and supports the following search for these documents. The process of automatically annotating solution documents using the ontology and its evaluation is described in detail in [12].

This paper focuses on improving the process of retrieving documents for similar technical solutions to overcome this specific barrier mentioned above. For this purpose, the general search for technical solutions provided by the existing ontology is enhanced by possibilities to offer not only exact search results for a certain user query, but enabling the search for similar solutions. As most relevant aspect for the similarity of technical solutions in automation industry, the shape of the object is proposed and discussed in this paper. Therefore, abstraction levels for the shape of objects in technical functions are developed and consequently added to the ontology. For illustration of the presented approach, publicly available documents in the automation industry serve as exemplary solution space. In these documents, companies provide information about previously installed solutions (e.g. a bottling and filling line for beverages). They are mostly used for marketing purposes to give references of previous work and are useful in generating first ideas how to approach an engineering task.

The paper is organized as follows: First, we will provide a short overview over the ontology and its use in the developed prototype. The technical functionalities of the prototype will not be described in detail and only as far as it is relevant for this work. Second, we describe the background of our approach for providing search for similar solutions by different abstraction levels. This paper will focus on the abstraction of the shape parameters of the handled objects. Then, we will describe the application of the search for similar solutions according to similar objects and show results of an initial test. We will conclude the paper with a discussion and summary of our findings and provide an outlook on the next steps to take.

2 DESCRIPTION OF THE ONTOLOGY AND ITS USAGE IN THE PROTOTYPE

2.1 Ontology for capturing solution knowledge
The developed ontology consists of concepts and relations needed for describing technical solutions. As explained in [9], the ontology was developed in an recursive process of textual analysis of existing solution documents in the field of automation industry and extraction of the most important concepts and relations. It is modeled as an OWL DL ontology using the Protégé ontology editor [13]. The base structure of this ontology with the core concepts is shown in Figure 2. In the ontology, the function has the central position. The function is realized by a technical solution. It is used in a special industrial sector, executed by a function owner, and performs a certain operation on a decent object. The more abstract concepts operation goal and property of an object serves for a more detailed description of the
operation and object of a function. Existing solutions can be described by instantiating these concepts with the appropriate instances.

![Diagram of the domain-specific ontology](image)

**Figure 2: Base structure of the domain-specific ontology (according to [9])**

### 2.2 Prototype for solution document annotation and retrieval

A prototype was implemented that uses the ontology to support the handling of knowledge about technical solutions. According to the general goal of storing, structuring and consequently retrieving knowledge, two major process steps are covered by the functionality of the prototype: the automatic annotation of textual solution documents and the subsequent search for these documents. For the automated annotation, the ontology serves as a vocabulary and provides the needed information about the existing relations of elements belonging to technical solutions. Therefore, the prototype uses the label property of the instances in the ontology to recognize the appropriate words and attach the corresponding concept and instance to the document. Linguistic algorithms analyze the sentences of a document and support in achieving a good annotations. For the search for solution documents, the ontology models the required concepts and relations of the technical solutions. According to the modeled dependencies in the ontology, similarities between technical solutions can be calculated (see section 3.1 for details).

The left side of Figure 3 shows a screenshot of this prototype with an exemplary result of an automated annotation. On the left side of the canvas, the annotated instances of a document are listed according to the concepts available as annotation filter (property of a solution, industrial sector, function, company and function owner). The annotated instances are added as metadata to the document. On the right side of the canvas, a graph browser offers the possibility to navigate through the existing ontology (instances and their relations) and adding further annotations manually. On the right side of Figure 3, a screenshot during the search process is shown. A query can be typed by the user and is interpreted by the prototype according to the modeled instances in the ontology. Recognized instances are interpreted as filter facts and used for filtering the search results. The results can be displayed in a list as well as in a graph browser.

![Screenshots of the prototype](image)

**Figure 3: Screenshots of the prototype during the document annotation (left) and the solution retrieval (right) process**
3 SIMILARITY OF TECHNICAL SOLUTIONS

In the first section of this chapter (3.1), an overview about the different general possibilities of similarity between technical solutions according to the information modeled in the ontology is given and demonstrated in an example with the prototype (3.2). Based on previous work and tests with the prototype, two promising types of similarity were identified in the current scenario of the automation industry: the similarity according to the goal of an operation (as described in [9]) and according to the shape of the handled objects. The similarity according to the shape of the handled objects was judged as very useful as most technical solutions in automation industry perform some kind of handling functions (e.g. gripping, transporting, and so forth). Common practices for abstracting these objects do not meet the actual requirements, so a new classification has to be developed. On the basis of a short summary of general approaches for abstraction of objects in the field of engineering design, industrial handling, and computer aided engineering, the approach of this work for abstracting objects by their shape and use this abstraction for the search for similar solutions will be explained in section 3.3. An application and short evaluation of this approach will conclude this chapter.

3.1 Similarity of solutions according to the ontology

The concepts and relations modeled in the ontology can be used for calculating the similarity of different technical solutions. In literature, several approaches exist for determining the similarity of entities within an ontology. Mostly, measures like graph distance or vector similarity are used (among others [14-16]). In our approach, we use a combination of distance in a subgraph of the ontology and the vector similarity. Like shown in Figure 4, a solution can be described by a subgraph of the overall ontology. The subgraph is built by the instances and relations that are connected with the instance of the solution. In this case, the two solutions A and B realize two functions and are provided by one company each. Function 1 of solution A and function 3 of solution B both share the “object a”. This “object a” is the connection between the two subgraphs. Thus, the two solutions are similar according to this intersection.

![Figure 4: Similarity of solutions according to the ontology](image)

Obviously, the degree of similarity depends on the amount of shared instances within the subgraphs of the compared solutions. Also, the distance of the shared instances from the original solutions is important for rating the similarity. Sharing only an object (which has the distance 2 from the solution) would be ranked lower than sharing a function (which has the distance 1 from the solution). Considering all possible relations in the current ontology, 15 different kinds of similarity could be used for calculating the similarity (according to industrial sector, used function carriers, performed functions, related solution properties, etc.).

3.2 Search for similar solutions in the prototype

The process of searching for similar solutions in the prototype is shown in Figure 5. Starting with a user query (either single words or as one or several sentences), the engineer can search for solutions within the repository. With the help of the ontology and linguistic algorithms, the developed prototype analyses the query according to its content. Similar to a subgraph of the ontology for a specific solution, the prototype interprets the user query as a subgraph. Thereupon, the recognized instances in the query are displayed for the user as filter facts in the prototype. This subgraph is compared with the subgraphs of the solutions in the ontology. Finally, the results of the query are calculated according to the similarity of the query-subgraph and the subgraphs of the solutions stored in the ontology and displayed in the prototype.
In the present state, five kinds of similarity of technical solutions were chosen as most important for the user: First of all, the similarity according to the industrial sector, in which the solution is used, is regarded. This covers important environmental circumstances for the solution. For example, a solution that is applied for chemical industry might not be applicable in food industry, as food safety and security factors have to be considered. Furthermore, the used function carriers of a solution, the realized functions (operations and objects), the company providing the solution and solution properties like robustness, cost-efficiency or cycle time are used as similarity measures.

To enhance the current capability of the prototype to search for similar solutions, further abstraction levels have to be added. Therefore, the required concepts, relations and instances have to be modeled in the ontology. When adding further abstraction levels to the ontology, several requirements have to be considered. On the one hand, the requirements concern the traceability and comprehensibility for the user searching similar solutions. This is why the proposed classification has to be generally understandable and the resulting similarities have to be conveyed in a comprehensible manner. On the other hand, issues concerning the effort for maintaining and updating the needed concepts and instances for calculating the similarity have to be considered. Two general possibilities have to be considered here: if a classification which contains the needed instances and relations yet exists, it should be integrated into the ontology. This possibility keeps the required effort relatively low. Though, if no existing complete classification can be used, the needed instances have to be added to the ontology and assigned to the appropriate classification manually. To reduce the emerging effort the proposed classification should be not too large and the assignment should be simplified for the user.

### 3.3 Abstraction of an object by its shape

For the present use-case in the automation industry, the abstraction of an object by its physical shape seems to be most promising for improving the search for similar solutions. For illustrating the background of the proposed shape abstraction, this section shows existing approaches for classifying objects in different fields of engineering design. The covered range spans from general approaches over industry-sector-specific classifications to the basic of geometry and shape definition in computer-aided design (CAD) applications. In order to achieve the desired abstraction and therewith the improvement of the search for similar solutions with the prototype, the different approaches are evaluated according to the requirements in the present use-case.

Existing general approaches in engineering design classify objects by their physical properties in a very generic way (e.g. material, signal, energy [17]). These generic classifications seem to be applicable in the present scenario to a limited extend, as they only allow a very abstract classification that is not useful in the regarded solutions in automation industry. Beyond these general approaches, classifications in certain industrial sectors describe objects more specific. For example, a common classification in industrial handling exists, that describe the objects in a more detailed way by their respective physical manifestation [18] (see Figure 6).
This classification provides a more detailed distinction of physical objects and gives a good starting point for further detailing. As many objects handled in automation industry are piece goods, a more detailed classification is proposed in [18]. There, piece goods are classified according to their basic shape and behavior (sphere, bar, flat piece, etc.). Further definitions and classifications exist for special areas like packaging [19] or food packaging technology [20], but are not yet applied beyond these areas.

In addition to these industry-sector-specific classifications, research and applications in the field of CAD and CAE deal with similar problems. In this field, current approaches aim at reusing existing CAD data for reducing the modeling efforts. For this purpose, shape recognition and feature recognition tools and algorithms have been developed [21, 22]. Although these CAD-based tools and algorithms are not applicable for our approach, as only textual documents and no CAD data is available, the general ideas and methodologies can be useful. The developed approaches divide the CAD parts into smaller subparts and features and calculate the similarity of the parts by comparing these subparts and features. In the present use-case, information about feature descriptions (like threads or holes of objects) is mostly not available in the solution documents. Therefore, feature information is not applicable for the search for similar solution documents. However, for adaptation in the present use-case, the breakdown of a part into its subparts seems to be promising. In CAD theory, these subparts are often standardized primitives. For example, in constructive solid geometry modeling (CSG), geometric 3-D primitives are used for modeling CAD parts. As primitives, blocks (i.e., cube), prisms, spheres, cylinders, cones and torus are used. By the combination of these primitives with Boolean operators, complex surfaces of a CAD part can be created. These primitives are quite similar to the classification of basic shape in industrial handling proposed by [18]. Therefore, they are chosen and used as initial point for a more detailed shape classification.

The following Table 1 shows the results of the chosen classification for the present application. The name of the shape is given with the according origin and a description of the meaning or an example. Eight shapes have their origin in either CSG modeling or the classification of industrial handling and were found useful for this classification.

<table>
<thead>
<tr>
<th>Shape</th>
<th>Origin</th>
<th>Description/Example</th>
<th>Shape</th>
<th>Origin</th>
<th>Description/Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>block</td>
<td>CSG / industrial handling</td>
<td><img src="Image" alt="Block" /></td>
<td>cone</td>
<td>CSG / industrial handling</td>
<td><img src="Image" alt="Cone" /></td>
</tr>
<tr>
<td>chock</td>
<td>CSG / industrial handling</td>
<td><img src="Image" alt="Chock" /></td>
<td>pyramid</td>
<td>CSG / industrial handling</td>
<td><img src="Image" alt="Pyramid" /></td>
</tr>
<tr>
<td>prism</td>
<td>CSG</td>
<td><img src="Image" alt="Prism" /></td>
<td>sphere</td>
<td>CSG / industrial handling</td>
<td><img src="Image" alt="Sphere" /></td>
</tr>
<tr>
<td>cylinder</td>
<td>CSG / industrial handling</td>
<td><img src="Image" alt="Cylinder" /></td>
<td>torus</td>
<td>CSG</td>
<td><img src="Image" alt="Torus" /></td>
</tr>
</tbody>
</table>
When testing the assignment of objects to these eight shapes, it became obvious that more shapes have to be added. Several objects could not be assigned to one of the existing shapes. To overcome this shortcoming, the following five shapes mentioned in Table 2 were added to close the existing gaps.

### Table 2: New classifications for the shape of objects

<table>
<thead>
<tr>
<th>Shape</th>
<th>Origin</th>
<th>Description/ Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-dimensional</td>
<td>New</td>
<td>Used, if the length is much greater than the and the width and the height of the object (e.g. a wooden stick)</td>
</tr>
<tr>
<td>2-dimensional</td>
<td>New</td>
<td>Used, if the length and the width much greater than the height of the object (e.g. a steel plate)</td>
</tr>
<tr>
<td>flexible</td>
<td>New</td>
<td>Used, if the object’s shape varies over time (e.g. an easily moldable plastic bag)</td>
</tr>
<tr>
<td>complex</td>
<td>New</td>
<td>Used, if the shape is too complex for assign to any of the shapes</td>
</tr>
<tr>
<td>not sure</td>
<td>New</td>
<td>Used, if there is no information about the shape of the object</td>
</tr>
</tbody>
</table>

### 3.4 Application in the use-case

The developed classification is tested by exemplarily assigning the objects stored in the ontology. A total of 100 objects from 30 different solution documents in the field of automation industry were assigned. The following Figure 7 shows an excerpt of this assignment. In the matrix, the objects of the ontology are mapped manually to the appropriate object shape. With this mapping, the objects are classified according to their shape. As shown in the figure, for example, a certain bottle and a certain can are both assigned to the shape “cylinder”. Therefore, solutions that handle a bottle or a can would be similar according to this similarity of the handled objects. As modeled in the matrix, ski poles are also assigned to the cylindrical shape. Thus, solutions that handle carrots or ski poles would be similar to these solutions, too; as both handle cylindrical objects. As this might be not desired, a more detailed distinction between objects is provided by the opportunity of assigning them to more than one shape. For example, the ski pole has a cylindrical shape, but a much greater length than its width or height, so it is additionally assigned to “1-dimensional”. With this multi-assignment, the similarity can be specified in more detail. Furthermore, it can be useful in case it is not clear what abstract shape to assign correctly or if one object can be assigned to different shapes. Considering this multi-assignment, the degree of similarity increases with the number of similar object shapes.

The following conclusions concerning the chosen shapes and the level of detail of the classification can be drawn by the actual assignment of the objects. All objects handled in the documents could be assigned to at least one shape. Most objects have either a block-shape (21%) or a cylinder-shape (28%). The chock-, prism-, cone- and pyramid-shapes were actually not used within the current set of
solution documents. Especially the 1-dimensional and 2-dimensional shapes were considered as very useful for the assignment. Furthermore, the quite small amount of different shapes keeps the manual effort for assigning the objects in acceptable boundaries.

The assigned objects and shapes were integrated into the ontology and initially tested on the search for similar solutions. Queries including the respective objects were tested on the set of solution documents and the results evaluated according to the usefulness of the solution. Two main insights have to be mentioned concerning the reaction of the test persons: first of all, the general approach seems to be correct as the test persons understood the chosen shape abstractions. Furthermore, the usefulness of the search for similar solutions was mostly evaluated positive, in particular if the similarity of the relevant object was comprehensible for the user. In some cases, the assignment of a certain shape to an object was questioned, as the test person did not understand the choice of this particular shape, and would rather have chosen a different shape. As a possible improvement for this understanding, the visual explanation of the shapes with a little symbol or an example was mentioned. Additionally, the high amount of the returned solutions has to be considered carefully as the number of possibly suitable solutions is increasing due to the fact that similar solutions are now presented additionally.

4. RELATED WORK AND DISCUSSION

In engineering design research, several approaches exist for capturing, storing and providing knowledge to support an engineer’s work. Several of them use ontologies for modeling the relevant knowledge (for an overview, see [8]). For example, Li et al. use an ontology where several information resources are combined for structuring engineering knowledge and supporting information retrieval in product development [23]. The shape features in their ontology are used for describing “devices” (objects) in a very detailed way by integrating existing taxonomies like STEP AP224 and vocabularies of major CAD packages. Ahmed et al. developed the EDIT-ontology (Engineering Design Integrated Taxonomy) for providing a visible indexing structure to users searching for knowledge [24]. Within these developed ontologies, a classification of the knowledge (stored in textual documents, reports, etc.) according to the concepts and relations can be realized. As every classification causes a similarity concerning the chosen concepts, these ontologies can also be used for the calculation of similarities between different entities (concepts or instances). Despite this capability, these approaches do neither focus on search for similar technical solutions nor on modeling similarities of objects in particular.

The findings of this work can be used not only in the field of automation industry. The general approach for retrieving similar documents can be applied in every knowledge-intensive area and only depends on the modeled concepts and instances in the ontology. When applying the approach in other areas of mechanical engineering (for example automotive industry), it has to be assured that the objects handled have the same meaning in the respective domain. Especially words that have a different meaning in different contexts (homonyms) have to be eliminated as they would lead to unrealistic associations. A current limitation of the presented approach lies in the fact that the assignment of the objects to their appropriate shape has to be done manually. Beside the resulting amount of work, errors made by the person assigning the objects reduce the quality of the search results. But as the emerging object-shape repository is increasing with every object added, this disadvantage will decrease over time.

5. CONCLUSION AND OUTLOOK

The abstraction of objects by the shape done in this research permits to improve the search for similar technical solutions in the field of automation industry. To achieve an adequate classification for this shape abstraction, different approaches for the classification of objects were combined and the resulting relevant concepts and relations added to the existing ontology. The chosen classification was successfully initially tested and evaluated. During the tests, possibilities for improvements and further development were identified. For example, the comprehensibility of the assigned abstract shapes can be overcome by providing examples of the corresponding shapes to the persons interpreting the resulting similarity.

Further work will aim at improving the identified shortcomings and cover additional testing of the presented approach. Issues will be regarded like the search within a larger solution space with more solution documents. Then, the implementation of user-interactive functionalities for reducing the possibly high amount of returned similar solutions like step-by-step reduction of the resulting solution
space seems to be promising. While in the present use-case within the automation industry, the goal of an operation and the shape of the handled objects are focused, other types of similarity can be relevant in other areas. Therefore, further abstraction levels will be added to the ontology for increasing the possibilities of searching similar solutions. When combining several of these similarity criteria, adequate means of user integration have to be considered for guaranteeing a comprehensible application for retrieving similar solutions. For example, personalized and step-wise enabling or disabling of different similarity-options and proper visualization techniques are planned to be integrated in the developed prototype.

ACKNOWLEDGEMENTS
This work has been funded by the German Federal Ministry of Economy and Technology (BMWi) through THESEUS. The authors wish to acknowledge gratitude and appreciation to all the PROCESSUS project partners for their contribution during the development of various ideas and concepts presented in this paper.

REFERENCES


Andreas Kohn, Dipl.-Ing.
Technische Universität München
Institute of Product Development
Boltzmannstr. 15
D-85748 Garching
Germany
Tel:+49 89 289 15124
Fax: +49 89 289 15144
Email: andreas.kohn@pe.mw.tum.de
URL: www.pe.mw.tum.de

Dipl.-Ing. Andreas Kohn works as a research assistant at the Institute of Product Development at the Technische Universität München, Germany. His scientific research covers computational intelligence for information retrieval in mechanical engineering and applied methods for complexity management in the field of product development.

Udo Lindemann is a full professor at the Technische Universität München, Germany, and has been the head of the Institute of Product Development since 1995, having published several books and papers on engineering design. He is committed in multiple institutions, among others as Vice President of the Design Society and as an active member of the German Academy of Science and Engineering.