Journal of Universal Computer Science, vol. 23, no. 9 (2017), 932-951
submitted: 31/7/17, accepted: 26/9/17, appeared: 28/9/17

Contextualization and Recommendation of Annotations to Enhance Information Exchange in Assembly Assistance[1]

Rebekka Alm
(Universität Rostock, Germany
Fraunhofer IGD, Rostock, Germany
rebekka.alm@uni-rostock.de)

Abstract: Increasingly flexible production processes require intelligent assistance systems containing information and knowledge to maintain high quality and efficiency. To ensure a reliable supply of information, it is of great importance to find easy and fast ways to record and store "new" information, as well as to provide a sensible mechanism to supply the information when needed.

In this paper an approach is presented that uses annotations in combination with a formalized knowledge base that represents the work domain. This pre-condition enables a context-based annotation recommendation. A framework is proposed to integrate different factors to measure the relevance of an annotation according to a given situation. The approach is illustrated using the example of an assembly assistance system.

To evaluate the users' attitude regarding annotations as instruction support and to test the system's capabilities when handling a great number of annotations some studies were performed and analyzed.
Key Words: metadata, knowledge management, assembly assistance
Category: H.4.3, H.5, H.3

1 Introduction

To cope with the increasing mass customization in production, former rigid product processes are substituted by more flexible yet also more specialized processes. To maintain equally high quality and efficiency despite the increased flexibility "information and knowledge are the firm's strategically most important resources today" [Widen-Wurff 2014]. The knowledge an experienced worker has acquired over time is in this regard particularly valuable.

Motivation and communication barriers are still a great obstacle to sharing knowledge [Connelly et al. 2014], even as the importance of information sharing is already widely accepted [McInerney 2002, Wang et al. 2004]. Therefore, it is of great importance to find an easy and fast way to record and store "new" information, as well as to provide a sensible mechanism to access the information when needed.

[1] This is an extended version of the paper Facilitating information exchange in assembly assistance by recommending contextualized annotations, presented at the First Workshop on Recommender Systems and Big Data Analytics co-located with I-KNOW'2016 in Graz, Austria, October 2016.

We propose to use content annotations as an intuitive interaction means to digitally communicate context-related information. We understand annotations as content objects (e.g. text snippets, photos) containing additional information about a related entity.

To enable further contextualization and re-usability of the annotations, we propose to formalize the information domain by an ontology enabling its concepts and instances in turn to be annotated with additional content. The unstructured annotation data is contextualized by linking it to the concepts and instances of the ontology modelling the domain context. The ontology in turn enables context-based recommendation of interesting annotations, that is the information is not only automatically provided in the exact context that it was created in, but also in another similar or related context.

In [Alm et al. 2015b] we proposed the usage of ontology-based annotations as missing link between the tacit knowledge of a worker and an intelligent assistance system. We showed the deeper integration of conceptual knowledge modeled in ontology-based annotations with procedural knowledge in cognitive architectures. This paper considers further contextual factors that influence the information's relevance and can be used to automatically recommend interesting annotations. We examine how annotations can be utilized as intuitive and easy means to contextualize content in an intelligent information system and propose measurements to quantify an information's relevance to a given situation. After introducing annotations as means for information communication in general, we present related work, before getting in detail for our use case of the work domain. We explain the architecture enabling the presented features briefly and explain our information selection framework integrating different contextual factors. A study evaluates factors influencing the user's acceptance of the digital content annotations and the overall performance of the system. This paper is an modified version of [Alm and Urban 2016] which focused more on the details of the selection algorithm and contained no study.

2 Annotations

Annotations can be found anywhere where people want to communicate a missing or new information, e.g. on notice boards or as placard on a wall [Coiera 2014]. With the rise of the Internet and digital connectivity, information resources are shared in groups. This opens digital possibilities for collectively discussing and enriching documents by digital annotations.

A digital annotation is in this context an additional content that relates to an already existing content, increasing this existing content by providing an additional layer of explanation [Agosti and Ferro 2007]. The annotation approach supports both communication sides in an intuitive way: annotations are easily created as well as accessed. Where complicated and strict forms or input masks confuse and demotivate, the annotation approach scores with the freedom to comment a content quickly and easily.

3 Related Work

Besides looking at "traditional" collaborative annotation systems, we take a look at relevant technologies from other areas that tackle challenges of information communication. As communication consists of two parts, recipient and transmitter, relevant technologies address on the one hand context-aware information provision , while on the other hand the transmitting side addresses technologies to integrate and contextualize the information.

3.1 Collaborative Annotation Systems

Collaborative annotation systems usually enable users to enrich documents with additional content, such as comments. The user can browse the document's content while getting additional information created by others as marginal notes or similar to sticky notes on top of the original content. Famous examples are the collaboration features of MS Word or Adobe Acrobat Reader allowing the user to annotate documents. The user does not have to search for information, the annotation is usually visible where it belongs to or at least a marker is seen to point out the presence of an annotation. Web-based collaborative visual analysis systems, such as sens.us [Heer et al. 2009] and Many Eyes [Viegas et al. 2007], support collaboration by allowing analysts to link text comments and graphic annotations to specific views or states of an interactive visualization. Other approaches mimic a board with post-it notes for brainstorming purposes, e.g. Wright et al.'s Sandbox [Wright et al. 2006].

All those systems have in common, that they do not handle a large number of annotations well. The visibility of the annotations or the original content as well as the "cognitive load" can be impaired, as the annotations are usually not filtered. Furthermore, the regarded "context" is very limited to the annotated content.

3.2 Context-aware Information Provision

Traditional knowledge management systems rely on the user to explicitly search for additional information by himself due to a demand regarding his current task [Wang et al. 2004]. The field of information retrieval researches methods to find information according to requested search terms. However, as the users require a more easy and intuitive means that minimizes this additional effort to get to the information, we are focusing on ways that automatically provide relevant information.

Recommender systems are a subclass of the information filtering systems that attempt to predict a “rating” or “preference” that a user would assign to an item. Based on this rating, a recommender system provides information to the user. Popular domains of research are movie, shopping, document and book recommendations [Park et al. 2012]. The traditional recommender systems usually disregard the notion of “situated actions” [Suchman 1987], the fact that users interact with the system within a particular “context” and that preferences for items within one context may be different from those in another context. Context-aware recommender systems [Adomavicius and Tuzhilin 2011] define a context in order to create more intelligent and useful recommendations. Cantador and Castells [Cantador and Castells 2009] for example propose semantic contextualization for their news recommender system news@hand.

Appropriate recommendations facilitate the discovery of relevant and interesting objects for the user. Yet, recommender systems rarely provide intuitive and simple ways to create and integrate new information. For our work, we make use of the simplicity of the annotations, but enrich them with ranking mechanisms to enable our system to recommend contextual relevant information to the user.

3.3 Integration and Contextualization of Information

Content annotations get some form of “context” by their reference to the annotated content, but there is usually no further contextualization. Thus, the annotations are well integrated in the annotated object, but the lack of indexing and further contextualizing results in a lack of possibilities for re-use or retrieval of the annotation’s content [Lortal et al. 2005].

A popular approach for contextualization of information is the enriching of documents with meta-data by associating ontology-concepts. Research in this field is summarized as “semantic annotation” [Uren et al. 2006]. The information’s context is in many research approaches formalized in an ontology. The greatest challenge of semantic annotation is the (semi-)automatic (and right) annotation of the information object as manual contextualization is seen as too time-consuming [Rodríguez-García et al. 2014].

Our work uses on the one hand content annotations for integration of additional information and on the other hand a domain ontology for further contextualization of the information similar to the functionality of semantic annotation.

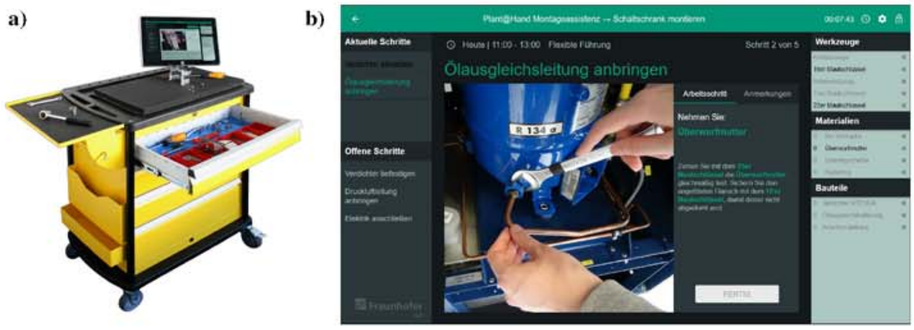

Figure 1: The Plant@Hand assembly assistant by Fraunhofer IGD: a) An assembly trolley is equipped with an assistance system [Alm et al. 2015b]. b) The display at the work station shows information supporting the current work task.

4 Information Communication using Annotations

Our objective is to provide the user with a quick and easy means to access and capture additional information. The digitization of the annotation activity enables several new possibilities for utilizing them as communication means. Since the annotations represent an own content, it is conceivable that they are not only of interest in reference to a single object. Their content can describe more than one object or give enriching information about groups of objects.

4.1 Use Case: Assembly Assistance

In production, there is an increasing demand for individual products, so that small batch sizes and flexible work processes are required. To ensure the quality of the work, the manual assembly should be supported by digital means. An assistance system can display information on the work process to optimize and control the process. User-authored annotations complement the work instructions.

Workers gather in the course of their work experiences that are useful when implementing tasks. We want to offer the opportunity to explicate the implicit knowledge of more experienced workers. If a person generates annotations explaining something noteworthy, everyone using the annotation system can benefit from the information. Hereby, the digitization of the annotation activity facilitates the re-usability of annotations and allows for mechanisms enabling their (context-based) retrieval.

For a better understanding we want to give the following example: A worker is supported in his task by a system that provides them with information regarding the current work task such as by step-by-step instructions and manuals. An example of such an system is the Plant@Hand Assembly Assistant shown in Figure 1. As the worker is experienced with a specific material used in the current work task, we consider it very desirable that they explicit their knowledge regarding specifics of this material (e.g. "be careful with tools as the surface of this material scratches very easily"). The additional information is swiftly and easily added as annotation into the assistance system. As the material is used in different tasks, the information regarding this material is also of interest to different tasks than the one that resulted in the creation of the annotation. Therefore, this added information should be provided in the context of each task which is using the same material. As the user does not know of the existence of an interesting annotation, it cannot be assumed that he will search for it. So we automatically recommend relevant and interesting annotations to the user. This way, the user is well informed and additionally can discover new and interesting information provided by colleagues.

We utilize user-authored annotations and the "sticky-note" metaphor as fast and intuitive means to explicate tacit knowledge. As a consequence we encounter the challenge to (semi-) automatically contextualize the annotations to use them in different contexts. We have to find a way to measure the "relevance" of an annotation regarding a specific context to determine if it should be recommended.

4.2 Architecture Solution

A first design draft of our architecture was presented in [Alm et al. 2015a]. We use a domain ontology modeling real-world objects and their relationships to formalize the information context. Real-world objects are in this context not necessarily only physical objects, but also abstract concepts such as a work task and its steps. From this ontology only that information has to be extracted that is relevant for a user's current situation. We utilize information retrieval mechanisms working on the structure of the ontology and the annotations attributes for this context-based information extraction. Figure 2 shows an overview of the concept. In summary, the following topics have to be considered:

1) **Capturing new information:** Establishing methods for the intuitive and unobtrusive collection of information during the work process.
2) **Organizing information:** Determining and modelling the information's context.
3) **Managing information:** Storing and accessing the information.
4) **Selecting information:** Ranking and selecting information in consideration of the information demand of the user's situation.
5) **Delivering information:** Providing the selected information by context-sensitive annotations.

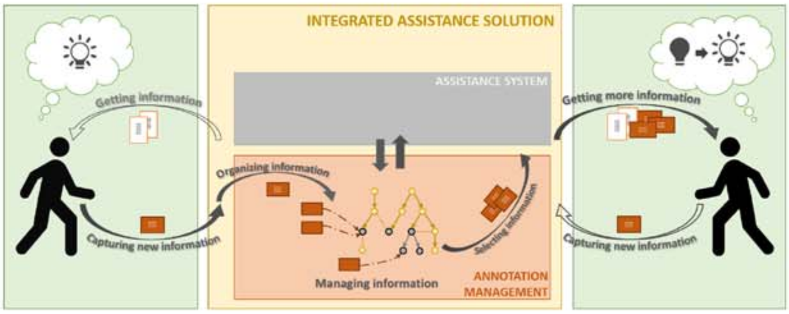

Figure 2: Concept Overview. A knowledgeable worker (left side) gets information from the system such as instructions. The worker knows more than the system and adds his comments. These get semantically enriched by topical classification that is deducted from the worker's situation. Regarding another (less knowledgeable) worker (right side) in another task the system selects and ranks topical relevant information and provides it to the worker to support him in his task and in getting more knowledgeable.

Those topics emerge to a circle of information sharing between humans. We demonstrated how the annotation activity can help to easily and intuitively insert information into the system. In the following we want to explain how we identify the relevant information.

4.3 Information Ranking and Selection by Contextual Factors

If unimportant or too much information is shown to a person, they are challenged to search for the relevant information themselves. The necessary additional effort causes the user to be fatigued and generally demotivated to browse the information. We use different contextual factors to measure the information's relevance to a given situation. We utilize a number of measures and combine them to an aggregated information selection measure.

4.3.1 Contextual Factors

Ontology Distance. Our first measure quantifies the relevance of the annotated ontology concepts to our current situation utilizing the ontology's graph structure. The ontology formalizes tasks and related things of a specific domain. Based on the concept c_1 of the ontology representing the current task, a concept c_2 of the ontology is considered to suggest relevant information if it is "near" the concept c_1, i.e. the measured distance between c_1 and c_2 is below a limiting value.

A $path(c_1, c_2)$ consists of a starting concept c_1 following different kind of relations to an ending concept c_2. The edges of a single relationship path $path_X(c_1, c_2)$ are all of the same type X. To measure the distance between c_1 and c_2 we have to differentiate between paths containing upwards and downwards directed (hierarchical) relationships ($X \in \{isA, partOf, ...\}$), paths containing horizontal (non-hierarchical) relationships ($X \in \{uses, ...\}$), and paths containing hierarchical and non-hierarchical (mixed) relationships. For **hierarchical** paths we use Resnik's [Resnik 1995] notion of a concept's information content IC. We utilize the method that Mazuel and Sabouret [Mazuel and Sabouret 2008] derived from the measure that was proposed by Jiang and Conrath [Jiang and Conrath 1997]:

$$W(path_{X \in \{isA, partOf, ...\}}(c_1, c_2)) = |IC(c_1) - IC(c_2)| \tag{1}$$

For the **non-hierarchical** relationships a different weighting approach is needed. Mazuel and Sabouret [Mazuel and Sabouret 2008] associate an individual weight TC_X for each relationship type X that is non-hierarchical:

$$W(path_{X \in \{uses, ...\}}(c_1, c_2)) = TC_X \times \frac{|path_X(c_1, c_2)|}{|path_X(c_1, c_2)| + 1} \tag{2}$$

Hereby, TC_X represents the semantic cost of the relationship type X.

To weight the distance between the ontology concepts c_1 and c_2 connected by **mixed relationships** the $path(c_1, c_2)$ has to be factorized into an ordered set $F(path(c_1, c_2))$ of n single-relation sub-paths. The distance between c_1 and c_2 is the sum of the weights of the sub-paths of the factorization $F(path(c_1, c_2))$. If there are multiple factorizations possible, the factorization resulting in the lowest W is preferred. In a previous work [Alm and Hadlak 2015] we focused on this relevance algorithm working a distance measure on the ontology. However, this path weight measure only gives a first ranking of which annotations are fitting for a given task. We found that additionally to this relevance measure, it is sensible to consider other factors for further ranking and filtering of the annotations. Therefore we consider further criteria to evaluate the quality of the annotations especially regarding their being helpful in supporting and educating the worker during his task.

Time. Hereby, time is an interesting factor, as an "old" annotation can be out-of-date while a "new" one has great potential to be more interesting. However, older annotations can also still be useful if their content is time independent. Therefore the "age" alone of the annotation is no reliable indicator of the annotations quality. Jan et al. [Jan et al. 2015] propose that useful annotations are reread. They define two quantity factors T_1 and T_2 where T_1 indicates the duration since the annotation was last visited and T_2 describes the average time between successive readings of an annotation. T_1 and T_2 accept values greater than 0, where a smaller value of T_1 and T_2 respectively indicates a higher quality of the annotation regarding their helpfulness.

Viewer Feedback. Furthermore, annotations can be rated by a viewer's feedback fb as helpful or useless. The feedback can be given explicitly by the user, giving him some option to vote. The feedback could also be derived automatically, where e.g. reading an annotation is regarded as positive feedback while ignoring an annotation is regarded as negative feedback. To use the individual reviews as a metric, the number of positive reviews and the number of negative reviews concerning a specific annotation are put in relation to each other. fb accepts values between -1 and 1 where a value close to 1 indicates a very positive regarded annotation while a value close to -1 indicates a very negative assessment of the annotation.

Prioritization. Additionally, a prioritization can be assigned to the annotation at creation time. The annotation's creator can state, if the annotation contains information that is necessary to prevent errors (high priority) or information to help improve the work process (normal priority). The priority $prio$ can be quantified by one value per annotation.

Expertise. The expertise of the annotation's creator affects the annotation's quality. If an expert has worked in the field for a long time, it is reasonable to assume that the knowledge they posses is valuable. Consequently, the annotations created by experts can be assumed to be especially valuable. As we presume that the elders will teach the younger employees and that the youngers will create corresponding annotations, the length of service is not necessarily a reliable factor. The annotation's quality improves if the annotation's creator has more experience with the annotation activity [Jan et al. 2015]. While the working expertise exp_w could be measured by the time a person has worked in the domain and can be taken from a personal profile, the annotation expertise exp_a of a $person$ could be measured by the number of annotations created and read by this $person$. Both measures exp_w and exp_a give values greater than 0 where a value near 0 marks a lack of experience. A higher value of $exp_{w/a}$ corresponds with a greater experience of the annotation's creator and thus indicates a "better" annotation.

4.3.2 Rank Aggregation

The different measures (W, T_1, T_2, fb, $prio$, exp_w, exp_a) quantify the annotations in different ways, so it is difficult to integrate them into a single measure. Rank fusion or aggregation is in information retrieval defined as the problem of combining a set of ranking lists in such a way to optimize the performance of the combination [Wei et al. 2010]. The traditional approach to integrate multiple ranking results from different individual rankers is to combine the ranking results (e.g., scores or ranks) produced by the individual rankers through certain rank aggregation techniques.

For combining our contextual factors, we first pre-select topically relevant annotations by the ontology distance W. Subsequently, we compile a set of several rankings

$$M = \{m_W, m_{T_1}, m_{T_2}, m_{fb}, m_{prio}, m_{exp_w}, m_{exp_a}\} \tag{3}$$

of this pre-selected set of annotations $A = \{a_1, ..., a_n\}$ using the presented measures. The resulting rankings are combined utilizing an easy and powerful approach from voting theory: the Borda count [van Erp et al. 2000]. The Borda count is an election method in which voters rank options in order of preference. The winner of an election is determined by giving each option, for each ballot, a number of points corresponding to the number of candidates ranked lower. Once all votes have been counted the option with the most points is the winner. In our case, the different rankings are the voters that sort the annotation options by preference according to the applied measurement. A ranking $m \in M$ of an annotation a determines the place of this annotation in the ranking $m(a) = m : a \mapsto r$ where $a \in A$ and $r \in \{1, ..., n\}$. We assign Borda points to the annotations according to their places in a ranking following the formula $1/place$, i.e. the first ranked annotation gets one point, the second ranked gets $1/2$ point, the third ranked gets $1/3$ point and the last place of n annotations gets $1/n$ points. The Borda points $bc_m(a)$ assigned to each annotation $a \in A$ for each ranking $m \in M$ can be defined as:

$$bc_m(a) = \frac{1}{m(a)} \tag{4}$$

The total Borda count score of each annotation a is then the sum of the Borda points of this annotation for all rankings. The original Borda count treats all classifiers/voters equal. To indicate a different "importance" of the measures, we propose to use a multiplier TX_m to weight each measurement before summing the individual Borda count values:

$$BC_a = \sum\nolimits_{m \in M} TX_m bc_m(a) \tag{5}$$

For instance, the feedback measure can get a parameter of $TX_{fb} = 2$ to balance that the time and experience measures are already very influential because there are two measures representing time and experience values. A higher Borda count BC_a indicates a higher relevance of an annotation a to a given context. Accordingly, we sort and provide the annotations beginning with the one scoring the highest BC.

4.4 Implementation Example

The practical implementation of the assistance system is carried out in a smart assembly trolley as already seen in Figure 1. The Plant@Hand assembly trolley is equipped with an information provision system that is supported by a cognitive architecture. A monitor displays information that supports the current work process. The annotations complement the work instructions as (side) notes. There are two modes for annotating. The user can locate an annotation directly on a 3D model e.g. of an assembly part and write their comments. Secondly, you can create the annotation besides the work instructions.

Our example ontology models (amongst other tasks) the assembly of a compressor. Figure 3 gives an overview of this part of the ontology. It is defined what steps are part of the work task and especially what kind of materials (such as screws and nuts of different sizes) and tools (such as wrenches of different sizes) are needed for each work task step. Annotations are linked to the ontology concepts that represent the task and/or related things that are annotated.

When a worker is supported by our information system during the work task step of attaching a compressor, our recommendation tool selects interesting annotations that have a low weighted ontology distance to the "Verdichter befestigen" concept of the ontology in Figure 3. For the filtering we choose empirically a limiting value of $W < 3$ where the horizontal *uses* relation has a semantic cost of $TC_{uses} = 1$. There are 12 annotations in our example to fit this limit. The IDs of these selected annotations are summarized in the following list sorted by their value of W:

- W = 1 (1 horizontal link): 76, 78, 83, 84, 85, 86, 93, 94, 95, 96
- W = 1.79 (1 hierarchical link): 75
- W = 2.58 (1 horizontal + 1 hierarchical link): 101

Already with a limited number of annotations we see that there are many annotations with the same value of W as they have the same kind of relationship to the focused ontology concept. To still accomplish a sensible ranking of the annotations we apply some of the other introduced measures (W, T_1, T_2, fb, exp_w, exp_a) and rank them according to their aggregated Border Count BC. We applied a weight $TX_{fb} = 2$ for the feedback value while we left all other $TX = 1$ as we perceived the feedback measure as not influencing enough otherwise. The ranks each of the 12 annotations receives for the measures W, T_1, T_2, exp_a, exp_w, and fb, as well as the resulting BC score is summarized in the following Table 1.

The additional measures enable a ranking of the annotations with identical values of W. Furthermore, the annotations are assessed based on more than

ID	W	m_W	m_{T_1}	m_{T_2}	m_{exp_a}	m_{exp_w}	m_{fb}	BC
95	1	1	6	9	10	3	1	3.71
96	1	1	7	1	6	5	10	2.71
76	1	1	5	8	1	7	9	2.69
93	1	1	4	3	3	12	3	2.67
84	1	1	12	12	3	9	2	2.61
83	1	1	9	2	9	2	8	2.47
78	1	1	11	11	11	1	12	2.44
86	1	1	8	4	5	4	6	2.16
94	1	1	2	10	8	10	7	2.11
101	2.6	12	1	5	6	8	4	2.07
85	1	1	3	6	12	6	10	1.95
75	1.8	11	10	7	2	11	5	1.32

Table 1: The ranks each of the 12 pre-selected annotations receives for the different measures, sorted by the resulting BC score.

only one of their attributes. Because of this, the annotations "95" and "96" are favored even as they score badly in one aspect. Interestingly the annotation with the id "101" is even favored above annotation "85" which is nearer to the topic, but scores worse in the entirety of its attributes. This can be sensible to get recommendations of high quality annotations that score good in more than one aspect. The beforehand filtering of the annotations by their W value ensures that only topic relevant annotations are chosen as initial set. For the further sorting we consider the other contextual parameters to be equally important to find high quality annotations. If the distance to the topic is seen as much more important the method can be adapted to weight the distance more by using a

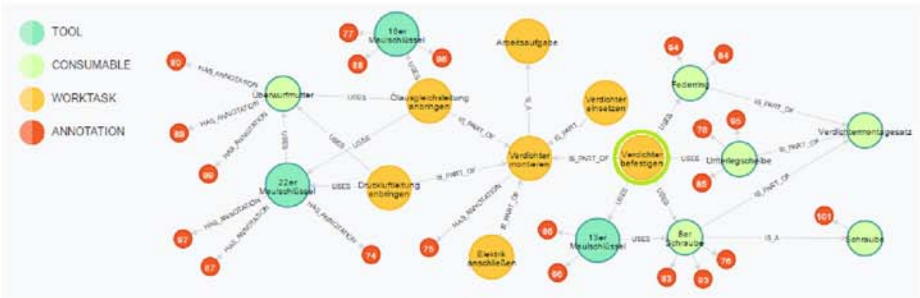

Figure 3: Part of our assembly domain ontology consisting especially of work task concepts and materials and tools. Annotations contain additional information describing the real life counter parts of those ontology concepts. The concept used in our sample calculation is circled green.

$TX_W > 1$ or use the BC value only to sort annotations of the same W value.

5 Evaluation

Until now no partner from industry has been found to evaluate the system in real factory conditions. Therefore, some tests were carried out under laboratory conditions to conduct a first evaluation of the following hypotheses:

- **H1:** Additional information available as annotations can improve the assistance. The contents and properties of the annotations influence the “usefulness” of the annotation for a given task.
- **H2:** The task of creating annotations is perceived as easier and faster than putting information into forms. The user is more willing to share his knowledge if this is done quickly and easily.
- **H3:** The system enables the handling of a great number of annotations.

5.1 Evaluation of Annotation Usage

The first test scenario was constructed to evaluate the “helpfulness” of the annotations. Only if the annotations can improve the assistance observable they will be accepted by the users. As we did not have assembly experts available for our evaluation we created a more common scenario to evaluate the perception of annotations by 20 users from different backgrounds. We asked our participants to fold paper stripes to create a so called “Froebel star” using a step by step folding instruction (see Figure 4). The instruction was not completely new to most users, but all of them conceded to not being able to fold this star without an instruction. The original instruction was taken from a website [2] and adapted as step-by-step instruction. Another version was annotated to include pointers and further suggestions. One group of 10 participants used first the “blank” original instruction before folding a second star using the annotated instruction. Another group (the control group) used first the annotated instruction before using the “blank” one. The time was documented as well as problems and errors arising while following the instructions. Both groups were afterwards asked about their perception and assessment of the annotations and the annotations’ properties.

The participants worked much more fluently when having additional annotations to enhance the instructions. Without the annotations, the participants were generally slower. They sometimes had problems to interpret the instructions and had to think longer about the text or even had to redo some steps. Furthermore, they missed optimizing potential given by the annotations. Being

[2] http://www.kreativ-insel.de/Weihnachtliches/Tech_Frobelsterne/Anleitung_Frobelstern/anleitung_frobelstern.html

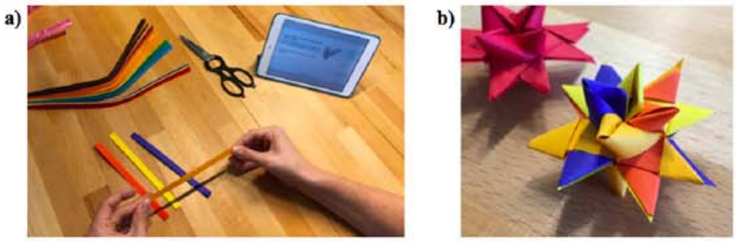

Figure 4: a) Study participants were asked to fold a "Froebel star" using instructions without and with annotations. b) Finished Froebel stars.

guided by the annotated instruction resulted overall in a better process time and end results. In some cases the original instruction was misinterpreted which resulted in faulty stars that could not be corrected. The annotations helped in learning the process with the result that the suggestions were still regarded when the following star was folded with "blank" instructions. The knowledge obtained by the annotations resulted in less frustration when following an instruction that was not well known to the users.

The perception of the annotations confirmed this positive impression. All study participants said it was the right amount of annotations. They assessed the annotations as "helpful" and "not obtrusive". When asked about the properties of the annotations they all ranked topic relevance as essential, while time properties, author experience and user rating were mostly assessed as important. Additionally most participants named comprehensibility as important property.

Overall, the first hypothesis **H1** could be confirmed. The annotations improved the process regarding the time as well as the result quality. Additionally the participants worked more content when being guided by the annotations while unguided participants were partly frustrated with the instruction. Consequently, the annotations contributed to a more happy work environment.

5.2 Evaluation of Annotation Creation

The second part of the test scenario was constructed to evaluate the convenience of creating annotations. If the users are able to insert their knowledge into the system fast and easily, the users will be more motivated to do so.

Our study participants from the previous scenario had used the existing annotations. Afterwards they were asked to create additional suggestions for the instruction using two different input methods. They created one annotation similar to a post-it note on the instruction and they were asked to fill an alternative form. Afterwards, they were asked about their perception and assessment of the creation methods.

The post-it note was created in less time. Filling the form required more time as the user had to read it first and comprehend what information is needed to fill each field. Therefore, it is not surprising that almost all users perceived the note as more pleasant. No one preferred the form, while only two participants declared to not care which way they used to insert new information.

All users said they are generally willing to insert information into the system and share their knowledge, but most of them also named some limitations. Most participants named time and stress as important factors . If they are very stressed to finish their work task in time they would not like to take additional time to interact with the system and include new information. A similar problem arises when the documentation activity is not considered part of the task, but has to be done in the user's "free time". When using the system in a work environment the superior has to support the annotation activity as something important and part of the work task. As further motivators for information sharing the participants named:

- Existing annotations: If others already created annotations, the users are more motivated to create annotations themselves.
- Positive feedback / prestige: If their past annotations are regarded as useful by other users they are motivated to share more information.
- Rewards: The commitment to support others with good advice is rewarded by the superior.

Altogether, the second hypothesis **H2** could be confirmed. Inserting information into the system was perceived and measured faster when done using annotations in comparison to forms. As time was named the most important (de-)motivator for information sharing, annotations can support a better information sharing. However, other aspects are still very important for the acceptance of the system. Especially the support of the superior was named one of the most important (de-)motivators.

5.3 Evaluation of Performance and Scalability

The last test scenario was constructed to evaluate the performance and scalibility of the system when handling a great amount of data. The system is only useful when it is able to run with a large and complex ontology and a great number of annotations. The cognitive load is taken from the user by the pre-selection of a certain set of fitting annotations, or at least facilitated by the ranking of the usefulness of the annotations. If the time constraint is also guaranteed, this would ensure that the system can handle a great amount of information in a sensible way.

For the test, the basic set was the ontology including annotations, which were already used in the implementation example of Section 4.4 for the calculation of the BC. The test ontology consists of 85 concepts, which are connected via 121 relationships. There are 35 annotations linked. Annotations were selected that complement the work task "Verdichter befestigen". The number of annotations was increased to 200 and 2,000 annotations respectively. In addition, the ontology was extended to provide an version of the ontology with 152 concepts and 251 relations as well as a further version with 303 concepts and 555 relationships. All three ontology sizes were tested with the same annotation sets that are located around the target "Verdichter befestigen". Figure 5 illustrates the test sets. The time was measured for the selection of annotations at run time. The one-time initialization time for the creation of the ontology has been ignored since this is no longer relevant for the running operation. The larger initialization time of a larger data set (because of larger ontology or more annotations) is also bearable because a longer time is acceptable during an initial installation. The calculations ran on a common workstation computer with Intel Core i5, 3.00GHz CPU, 16GB RAM and 64-bit operating system.

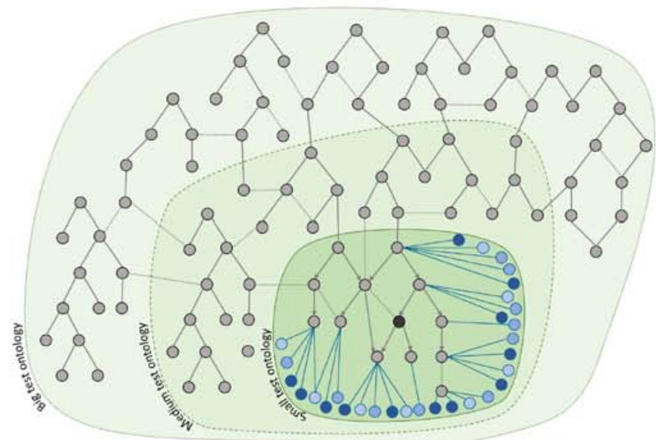

Figure 5: The "small test ontology" consists of concepts (gray dots) and connecting relationships (black lines) of the darkest area. There are annotations (blue dots) linked with blue lines that complement the focused work task "Verdichter befestigen" (darkest dot). The number of annotations considered was at first increased by including the lighter blue dots and then to include the lightest blue dots annotations. In addition, the ontology was extended to provide two extended versions of the ontology: "Medium test ontology" and "Big test ontology". All three ontology sizes were tested with the same annotation sets that are located in the darkest green area.

The performance test showed that the size of the ontology has nearly no effect on the system's run time performance of selecting fitting annotations to a given task. This is not unexpected as the algorithm focuses on a selected ontology concept and its related entities in a limited range. Parts of the ontology beyond this range are not relevant to the calculation. This is beneficial as ontologies might get very big and complex when modeling the context of work tasks and their related entities.

The number of annotations (in the considered range) is more influential for the calculation. The calculation of the basis set of 35 annotations needed circa 1.4 seconds when using our common computer. When increasing this number to 200 annotations the calculation consumed only little more time with 1.7 seconds. Even ten times of this number of annotations (2,000) still only needed about 3 seconds for calculation. All in all this is an acceptable frame of time to load annotations as the user will first need to conceive the work task instructions that are annotated before studying the additional information given by the annotations. When performing on a more powerful computer or server these times will be even lower. Furthermore, 2,000 seems like an incredible high number of annotations that are regarded for one focus task. Such a high number of additional remarks indicate something faulty. Possible reasons include:

– The original instruction could be erroneous or insufficient and should be corrected.

– The annotations could be redundant and should be combined or deleted.

– The work task could be modeled too complex including too many dependencies and links and should be split into further steps or specific cases.

Regardless of the sensibility of including so many annotations the performance test could confirmed the third hypothesis **H3**. The system can handle both a great number of annotations and a complex and big ontology. Table 2 gives an overview of the result times.

		Number of Annotations in DB		
		35	200	2000
Ontology size	85e/121r	$\approx 1.4s$	$\approx 1.7s$	$\approx 2.9s$
	152e/251r	$\approx 1.4s$	$\approx 1.7s$	$\approx 2.9s$
	303e/555r	$\approx 1.4s$	$\approx 1.7s$	$\approx 2.9s$

Table 2: The time needed to calculate the best annotations (by BC) for a given task compared by different data sizes. Ontology size is given by the number of entities (e) and relationships (r).

5.4 Study Summary and Conclusions

All three hypotheses could be confirmed by a two-part user study and the technical evaluation.

Regarding hypothesis **H1:** The additional information available as annotations did improve the assistance observably. The results were achieved faster and with better quality when the users were assisted by annotations. Furthermore, the users did feel less frustration and stress when being provided with the additional annotations. These arguments indicate that the annotations did indeed improve the assistance.

Regarding hypothesis **H2:** The task of creating annotations was perceived as easier and faster and thus more pleasant than putting information into forms. The users preferred the annotations over forms, but would still be reluctant to share their knowledge if it was not supported by their superior and colleagues.

Regarding hypothesis **H3:** The system does enable the handling of a great number of annotations. The technical evaluation demonstrated that the system works acceptable with a great and complex ontology as well as with a reasonable number of annotations.

All these findings have to be viewed critically because of the laboratory conditions of the test environment and must be reassured by field tests and with a bigger group of participants. Nevertheless these first findings show an approval of our general idea and motivation to improve assistance by annotations.

6 Summary

In this paper we introduced an approach for using annotations as an easy and intuitive means to capture new information and to recommend interesting annotations according to a given context. We enable a broader re-usability of the annotations by automatically recommending them to a user according to his current situation. Our recommendation mechanism includes a selection by a relatedness measurement of the annotations to the current context as well as a ranking of the annotations according to our assessment of several attributes. We showed how our method enabled a helpful information provision to an assembly work task. The approach is easily adaptable to other domains. Where human tasks are supported by an information system it is sensible to give the possibility to add missing or new information by annotations. The automatic recommendation of interesting annotations is also generally beneficial as users usually do not know what information to look for or are not motivated to search themselves for lack of time. The user study confirmed that recommended annotations not only help to keep the user well informed. They support the further education of the users by providing diverse information and especially by encouraging the exchange of experiences. Furthermore they enable especially inexperienced users to work more independent and with less frustration because of too vague instructions. Nonetheless is the acceptance of this approach also dependent on further aspects such as the support by superiors.

We intent to perform further experiments to explore which measures, parameters and weights are reasonable for specific use cases. Especially the scaling of the area of interest has to be researched in a real environment, i.e. which value of W is the limit. It can also be considered to enable the user to adjust this parameter to fit his information demand. Workers in training for example could be interested to get more annotations to enhance their learning process.

Acknowledgments

This research has been supported by the German Federal State of Mecklenburg-Western Pomerania and the European Social Fund under grant ESF/IV-BM-B35-0006/12.

References

[Adomavicius and Tuzhilin 2011] G. Adomavicius and A. Tuzhilin. Context-aware recommender systems. *Recommender systems handbook*, pages 67–80, 2011.

[Agosti and Ferro 2007] M. Agosti and N. Ferro. A formal model of annotations of digital content. *ACM Transactions on Information Systems*, 26(1):3–es, nov 2007.

[Alm et al. 2015a] R. Alm, M. Aehnelt, S. Hadlak, and B. Urban. *Human Interface and the Management of Information. Information and Knowledge Design: 17th International Conference, HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part I*, chapter Annotated Domain Ontologies for the Visualization of Heterogeneous Manufacturing Data, pages 3–14. Springer International Publishing, Cham, 2015.

[Alm et al. 2015b] R. Alm, M. Aehnelt, and B. Urban. Processing manufacturing knowledge with ontology-based annotations and cognitive architectures. In *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*, page 25. ACM, 2015.

[Alm and Hadlak 2015] R. Alm and S. Hadlak. Towards integration and management of contextualized information in the manufacturing environment by digital annotations. In *Proceedings of the International Summer School on Visual Computing 2015*, pages 141–157. Stuttgart: Fraunhofer Verlag, 2015.

[Alm and Urban 2016] R. Alm and B. Urban. Facilitating information exchange in assembly assistance by recommending contextualized annotations. In *Proceedings of the First Workshop on Recommender Systems and Big Data Analytics* co-located with I-KNOW'2016. Online under: http://socialcomputing.know-center.tugraz.at/rs-bda/papers/RS-BDA16_paper_2.pdf

[Cantador and Castells 2009] I. Cantador and P. Castells. Semantic Contextualisation in a News Recommender System. In *Workshop on Context-Aware Recommender Systems (CARS 2009)*, volume 1068, pages 19–25, 2009.

[Coiera 2014] E. Coiera. Communication spaces. *Journal of the American Medical Informatics Association : JAMIA*, 21(3):414–22, 2014.

[Connelly et al. 2014] C. E. Connelly, D. P. Ford, O. Turel, B. Gallupe, and D. Zweig. 'I'm busy (and competitive)!' Antecedents of knowledge sharing under pressure. *Knowledge Management Research & Practice*, 12(1):74–85, 2014.

[Heer et al. 2009] J. Heer, F. B. Viégas, and M. Wattenberg. Voyagers and Voyeurs: Supporting Asynchronous Collaborative Visualization. *Communications of the ACM*, 52(1):87–97, 2009.

[Jan et al. 2015] J.-C. Jan, C.-M. Chen, and P.-H. Huang. Enhancement of Digital Reading Performance by Using a Novel Web-based Collaborative Reading Annotation System with Two Quality Annotation Filtering Mechanisms. *International Journal of Human-Computer Studies*, 86:81–93, oct 2015.

[Jiang and Conrath 1997] J. Jiang and D. Conrath. Semantic similarity based on corpus statistics and lexical taxonomy. In *Proc. on International Conference on Research in Computational Linguistics*, pages 19–33, Taiwan, 1997.

[Lortal et al. 2005] G. Lortal, M. Lewkowicz, and A. Todirascu-Courtier. Annotation: textual media for cooperation. *IWAC*, 2005.

[Mazuel and Sabouret 2008] L. Mazuel and N. Sabouret. Semantic relatedness measure using object properties in an ontology. In A. Sheth et al. (eds.), *The Semantic Web - ISWC 2008*, volume 5318 of *Lecture Notes in Computer Science*, pages 681–694. Springer Berlin Heidelberg, 2008.

[McInerney 2002] C. McInerney. Knowledge management and the dynamic nature of knowledge. *Journal of the American Society for Information Science and Technology*, 53(12):1009–1018, oct 2002.

[Park et al. 2012] D. H. Park, H. K. Kim, I. Y. Choi, and J. K. Kim. A literature review and classification of recommender systems research. *Expert Systems with Applications*, 39(11):10059–10072, 2012.

[Resnik 1995] P. Resnik. Using information content to evaluate semantic similarity in a taxonomy. *Proc. of 14th Int. Joint Conf. on AI*, pages 445–453, 1995.

[Rodríguez-García et al. 2014] M. Á. Rodríguez-García, R. Valencia-García, F. García-Sánchez, and J. J. Samper-Zapater. Ontology-based annotation and retrieval of services in the cloud. *Knowledge-Based Systems*, 56:15–25, 2014.

[Suchman 1987] L. A. Suchman. *Plans and situated actions: the problem of human-machine communication.* Cambridge university press, 1987.

[Uren et al. 2006] V. Uren, P. Cimiano, J. Iria, S. Handschuh, M. Vargasvera, E. Motta, and F. Ciravegna. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Web Semantics: Science, Services and Agents on the World Wide Web*, 4(1):14–28, jan 2006.

[van Erp et al. 2000] M. van Erp and L. Schomaker. Variants Of The Borda Count Method For Combining Ranked Classifier Hypotheses. *Proceedings 7th International Workshop on frontiers in handwriting recognition*, pages 443–452, 2000.

[Viegas et al. 2007] F. B. Viegas, M. Wattenberg, F. Van Ham, J. Kriss, and M. McKeon. Many Eyes: A site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128, 2007.

[Wang et al. 2004] S. Wang, R. a. Noe, and Z.-M. Wang. Motivating Knowledge Sharing in Knowledge Management Systems: A Quasi-Field Experiment. *Journal of Management*, 40(4):978–1009, jul 2014.

[Wei et al. 2010] F. Wei, W. Li, and S. Liu. iRANK: A Rank-Learn-Combine Framework for Unsupervised Ensemble Ranking. *Journal of the American Society for Information Science and Technology*, 61(6):1232—-1243, 2010.

[Widen-Wurff 2014] G. Widen-Wulff. *The Challenges of Knowledge Sharing in Practice: A Social Approach.* Chandos Information Professional Series. Elsevier Science, 2014.

[Wright et al. 2006] W. Wright, D. Schroh, P. Proulx, A. Skaburskis, and B. Cort. The sandbox for analysis - Concepts and methods. *Conference on Human Factors in Computing Systems - Proceedings*, 2:801–810, 2006.