2013

Gathering Knowledge from Social Knowledge Management Environments: Validation of an Anticipatory Standard

René Peinl
Hof University of Applied Sciences, Hof, Germany, rene.peinl@hof-university.de

Lars Hetmank
TU Dresden, Dresden, Germany, lars.hetmank@tu-dresden.de

Markus Bick
ESCP Europe Business School Berlin, Berlin, Germany, mbick@escpeurope.eu

Stefan Thalmann
University of Innsbruck, Innsbruck, Austria, stefan.thalmann@uibk.ac.at

Paul Kruse
TU Dresden, Dresden, Germany, paul.kruse@tu-dresden.de

See next page for additional authors

Authors
René Peinl, Lars Hetmank, Markus Bick, Stefan Thalmann, Paul Kruse, Jan M. Pawlowski, Ronald Maier, and Isabella Seeber

Gathering Knowledge from Social Knowledge Management Environments: Validation of an Anticipatory Standard

René Peinl[1], Lars Hetmank[2], Markus Bick[3], Stefan Thalmann[4], Paul Kruse[2], Jan M. Pawlowski[5], Ronald Maier[4], and Isabella Seeber[4]

[1] Hof University of Applied Sciences, Hof, Germany
rene.peinl@hof-university.de
[2] TU Dresden, Dresden, Germany
{lars.hetmank,paul.kruse}@tu-dresden.de
[3] ESCP Europe Business School Berlin, Berlin, Germany
mbick@escpeurope.eu
[4] University of Innsbruck, Innsbruck, Austria
{stefan.thalmann,ronald.maier,isabella.seeber}@uibk.ac.at
[5] University of Jyväskylä, Jyväskylä, Finland
jan.pawlowski@jyu.fi

Abstract. Knowledge management is more and more happening in social environments, supported by social software. This directly changes the way knowledge workers interact and the way information and communication technology is used. Recent studies, striving to provide a more appropriate support for knowledge work, face challenges when eliciting knowledge from user activities and maintaining its situatedness in context. Corresponding solutions in such social environments are not interoperable due to a lack of appropriate standards. To bridge this gap, we propose and validate a first specification of an anticipatory standard in this field. We illustrate its application and utility analyzing three scenarios. As main result we analyze the lessons learned and provide insights into further research and development of our approach. By that we reach out to stimulate discussion and raise support for this initiative towards establishing standards in the domain of knowledge management.

Keywords: knowledge management, social software, standard, metadata

1 Introduction

Knowledge Management (KM) is more and more happening in social environments, supported by social computing [1], [2] and so-called social software (SSW) [3-5]. However, KM solutions in these environments are not interoperable due to a lack of appropriate standards. Thus, we propose and validate a specification as a proposal for a new standard and assess this in a practical environment. Standards are widely established in the information technology domain as an enabler of wide adoption and diffu-

11th International Conference on Wirtschaftsinformatik,
27th February – 01st March 2013, Leipzig, Germany

sion [6-8]. In the case of KM, only very few specific standards have been created and used so far, e.g., the European Standardization Body CEN [9]. When working with KM solutions in diverse social and SSW environments, interoperability becomes a key issue. However, interoperability is currently not represented well in such systems and does not exploit the opportunities for further supporting knowledge activities [10]. One major issue in this regard is the missing standardization [4] which hinders for example the exchange of contextual information and its usage within KM systems.

Therefore, we have proposed an ontology-based specification [11] as a starting point for an anticipatory standard [12] which allows to model, share and reuse different aspects relevant to KM such as context, activities and resources. It is important to notice that the proposal for an anticipatory standard is not a complete and final version but a validated starting point for a broad consensus process. This paper is the fourth in a series of papers striving to establish a standard for a knowledge container (KC) format that strives to join requirements stemming from a KM and Web 2.0 perspective. In our previous work we first identified the main challenges for KM in regard to SSW environments and consequently defined six central concepts [13]: comprising 1) knowledge activity (KA), a set of goal-directed actions within a user's context, 2) knowledge object (KO), a codified representation of externalized knowledge, 3) knowledge trace (KT), a representation of a user's action in a 4) knowledge activity stream (KAS), 5) knowledge bundle (KB), a collection of knowledge traces affiliated to a KO, and 6) knowledge container (KC), a compound document holding KO and KB on multiple aggregation levels. In the next step we further refined our understanding of a potential specification by setting the KC into relation with the creation of a semantic model [14]. As can be seen, we imply a pragmatic definition that sees knowledge as bound to people in general, but allows the term codified knowledge when speaking about contextualized information within documents [10]. In a next step, we reviewed existing standards to build on the work already done, added the knowledge worker (KW) as a further concept and proposed a first ontology for a KC [11].

The goal of this paper is to continue our prior work and to investigate how such an ontology-based standard could be implemented. We performed an initial validation from a technical point of view and gained insights from manually creating the technical artifacts envisioned by our KC. The insights gained comprise, e.g., limitations of current standards that we re-use, problems that can be foreseen when trying to derive the data automatically from the system instead of creating it manually as well as technical challenges with versioning of RDF data. The results of our first validation step will further aid in the constant improvement of the anticipatory specification and strives to stimulate discussions in our research communities.

After providing background on the key concepts used in this paper in section 2, the following sections 3 and 4 describe the adopted methodology and technical implementation, respectively, for this first validation step using an adaptation of the Reference Model Analysis Grid [15]. In section 6 we report on the results of this initial validation step as well as on lessons learned (LL).These insights provide starting points for future research and attempt to stimulate discussion for our initiative towards establishing a standard in the domain of KM.

2 Background

Information and communication technology (ICT) in general and SSW in particular [3] build the basis and means for knowledge workers to communicate and collaborate online [16] within or between teams or even beyond organizational boundaries. At the same time, SSW changes the way knowledge workers interact with each other and the way ICT is used [17]. Recent research studies, striving to provide a more appropriate support for knowledge work, face challenges when eliciting knowledge from user activities and maintaining its situatedness in context [18-20]. Unlike data and information, it is difficult to codify and transfer knowledge, because it highly depends on the social context of its creation [21]. SSW tools enable a richer capturing of context in which content has been produced, modified and used [4]. Additional features, such as activity streams [22], represent an example how context can be preserved to support knowledge workers during their knowledge activities. Nonetheless, knowledge workers still experience difficulties to identify the feed items that are of genuine interest [23] depending on the situation they are in. Major issue is the missing standardization [4] which hinders the exchange of contextual information and their usage within KM systems. In terms of the technological base feature, we are led by the notion of activity streams which consist of four components [23]: the subject, who carried out the action, the action, which caused some change in content or state, the object, on which the action was carried out, and the time, at which the action occurred [23]. We discuss the main concepts subject, action and object in the following:

Subject – The Knowledge Worker Perspective. A Knowledge worker is mainly characterized by ill-structured non-routine tasks in complex domains and often requires creative problem-solving abilities and specialized knowledge [23], [25]. The handling of non-routine tasks which are considered particularly important [19] requires that the knowledge worker is provided with IT tools that give at least partial or fragmented context information to enable the development of knowledge [26].

Object – The Knowledge Object Perspective. When conceptualizing knowledge as an object, it is situated and associated to the current context as well as with historical context [27]. The KO is the smallest explicit piece of documented knowledge [27] and needs to be described to be used in certain contexts [28]. By describing KOs with associated metadata [29], additional knowledge about the KO gets documented.

Action – The Knowledge Activity Perspective. In weakly structured work domains, it is difficult to automatically detect knowledge activities on a high level of granularity and subsequently to conclude from these high-level activities to low-level activities [19], [20], [30]. These low-level activities refer to easily traceable system operations such as to post, to bookmark etc. [31].

3 Methodology

This work follows a design science methodology and adopts the understanding that an artifact is complete and effective when it satisfies the requirements and constraints of the problem it was meant to solve [32]. According to Hevner et al. [32], scenarios can be applied as an evaluation technique for innovative artifacts, in particular for new and complex artifacts which cannot be evaluated as such in one step. "A scenario is a tool to explore a possible, plausible future by identifying key technical and social developments required for it to be realized" [33]. It aims at creating awareness of possible developments and have been used for new IS technologies with high social impact [34], [35]. Social KM environments arguably have a high social impact as this is at the core of their intended use and thus they are particularly suitable for such a scenario-based evaluation. In our initial validation step we use scenarios as a means to generate a number of problem domains to assess its applicability to real world practices (section 3.1). We adopt the Reference Model Analysis Grid (RMAG) which specifically aims at validating standards and suggests concrete measures for assessment [15] (section 3.2). As we are in the stage of pre-standardization, we focus on the economic and deployment perspective to anticipate the potential quality of the proposed standard. Consequently, the validation as presented does not evaluate the proposed solution completely. We selected RMAG perspectives that support our assessment at this stage without having the proposed ontology implemented yet.

3.1 Scenarios

We constructed three detailed scenarios around the technical implementation to demonstrate its utility as a proof-of-concept. The construction of such scenarios can already be seen as (first) descriptive evaluation [32]. All three were described upon typical application areas from real partner organizations of the authors containing main concepts we have introduced in our ontology [11].

Scenario 1: New Product Introduction. This scenario illustrates the activity of finding new product ideas. In the last strategy session of the company, Kurt, the project leader of the product development department, was instructed by the CEO to find new product ideas and to advance at least one of them as soon as possible towards a technical solution paper. Kurt does not want to waste any time to arrange a physical meeting and therefore decides to invite Elaine from the marketing department, Terry from the design department and Justin from the technical development department to a Skype conference. While Kurt is chairing the conference, he makes notes about the new product ideas and the next steps to be done by using the note-taking service Evernote. These steps include on the one hand a patent evaluation and on the other hand a market research. Kurt soon decides to use Google Docs to formalize both knowledge activities by creating a first draft of the technical solution paper and setting up an initial document structure. He also invites all project team members to add any related contributions to the document. Justin searches for relevant patents at Google's patent search engine at his home office at 2 a.m. in the morning. The next day, Justin creates a list of all patents that are held by the own or any foreign company by using Excel. Meanwhile, in order to better understand customer needs, Elaine conducts a quantitative market research. Therefore, she creates a standardized survey on the online survey platform FluidSurvey. A few days later, she exports several diagrams in png file format. Both Justin and Elaine add their contribution to the technical solution paper. Based on the diagram of the market survey, Terry from the design department makes a first sketch of the new product. The next day, Justin screens the current version of the technical solution paper and adds a comment to the sketch that claims some engineering issues that may occur. Terry sends an email with a URL link of the current version of the technical solution paper to April from the procurement department in order to discuss expenses related to material choice. After reading Justin's comment and receiving a detailed answer from April, Terry changes some aspects of the sketch. Finally, Justin creates a drawing that shows technical details of the new product and adds it to the technical solution paper. After Kurt returns from his business trip, he wants to get an overview of all recent events.

Scenario 2: Organization-spanning research project. This scenario illustrates the activity of developing an idea collaboratively to achieve a common goal. Bob has an interesting research idea and wants to find related literature on the Web. Due to the fact that the idea's topic is very new and innovative, no suitable journals, conference papers or books are available. However, during his search Bob identifies five interesting people working in this domain. First, he follows them on Twitter and re-tweets the most important messages via his twitter account. Additionally, he sets bookmarks on the blogs of these people that he deems by now to be experts in their respective domains. After following the activities of these five experts for a while, Bob decides to contact them to initiate a discussion about his research idea. For this purpose he creates a Google discussion group and writes some provocative questions related to the experts' expertise.

All five experts, Maria, Peter, Igor, Paul and Andreas, are highly interested in Bob's idea and an intensive discussion evolves. During the discussions the group decides to chat about Bob's research idea in a series of videoconferences. They exchange their interim results via e-mail, share links to Internet resources such as websites, recent conference and journal papers and publish interesting facts in their discussion group. Additionally they decide to write a position paper. Therefore, they store and share related literature, figures and other sketches in a shared DropBox folder. After the first presentation of the idea on a related workshop the group receives attention from a broader audience. The fast growing group of researchers working on Bob's idea makes it difficult to maintain an overview. Even though the core group around Bob still exists, the original idea has evolved. Increasingly the group experiences that due to their different knowledge they have on the same topic their discussions and information exchanges loses effectiveness which is why they decide write a further paper with the aim to structure and integrate their different viewpoints.

Scenario 3: IT consultancy. This scenario illustrates the activity of creating a software specification and project proposal based on a requirements document from the customer (tender).On Tuesday morning, the request for proposal (RfP) for an Intranet relaunch at a large pharmaceutical (BMS) arrives at Jessica's desk. She is a project manager for a small IT consultancy that specializes in Intranet and Internet projects based on open source software like Alfresco, Liferay and XWiki. She has seven working days to assemble a team and create a proposal including a rough technical specification of the offered solution. She therefore uses the VoIP server to call Marcus, the sales representative who already had two meetings with the customer to get a better understanding about the whole story. Then, she creates a new document workspace in the company's own Liferay installation, in order to collect all information she needs for the proposal and files the RfP there. After that, she creates a new proposal document from the company's template and fills in the formalities like customer name and date, stores it in the document workspace and creates a task for Jan, the software architect, to construct an overall architecture for the project. Jan receives the email notification about his new task and creates a technical drawing, showing the overall architecture in a Visio document. He inserts it in the proposal document and writes two pages of explanations about his ideas. Then he updates the task and creates two new ones for Julia, the Web designer of the company and Tom, one of the senior developers, to create a wireframe and screen design and to further detail the given work packages respectively. He further sets up a meeting to discuss the further procedure face to face. Julia creates her visuals with Inkscape and Gimp, inserts them into the proposal document and stores the source files in the document workspace. Tom writes an own chapter for each of the work packages, so that the proposal is about 15 pages long after he finished. Jan and Jessica both do a proof reading of the document and Jessica schedules a session with Jan and Tom in order to create a preliminary project plan and estimate the efforts. The final documents are then sent to Oliver, the CFO of the company, who discusses the efforts with Jessica and Marcus in order to fix a final price. After that, the proposal is sent to the customer by email.

3.2 Technical Implementation

Our anticipatory specification builds on existing standards as reviewed in [11]. Due to the intended semantic description, the contents should conform to RDF/XML and reference an OWL ontology we have created. We decided not to reuse vocabulary from different existing standards, but instead build our own vocabulary and create a mapping on the schema level. This can be done using either OWL equivalent classes or XSLT, if the model standard has no OWL schema, but operates on the XML level only. The creation of the ontology or schema level brought several challenges. The result can be seen in figure 1, which shows the classes, object properties (i.e. relations between classes) and data properties of our ontology.

One central element of the proposed ontology is the knowledge trace. The activitystrea.ms standard served as a role model. We adopted all attributes and entities and coded them using RDF/XML. Consequently actor, object and target of the KT are coded as separate resources and referenced from the KT. This creates locally redundant storages of entities which might hamper performance but should not further be considered at this point in time. KT are therefore structured similar to RDF itself: actor, verb, object instead of subject, predicate, object. The optional target specifies the system, container, or knowledge worker that represent the target of the action. Possible verbs and also object types are specified in the activitystrea.ms standard.

We have grouped similar properties under a common super property, e.g., knows as parent of colleague, follows and hasSupervisor or the abstract orgRelation which serves as parent for affiliation, memberOf and worksFor. This can be helpful for querying the data on a coarse level and helps compensating different interpretations.

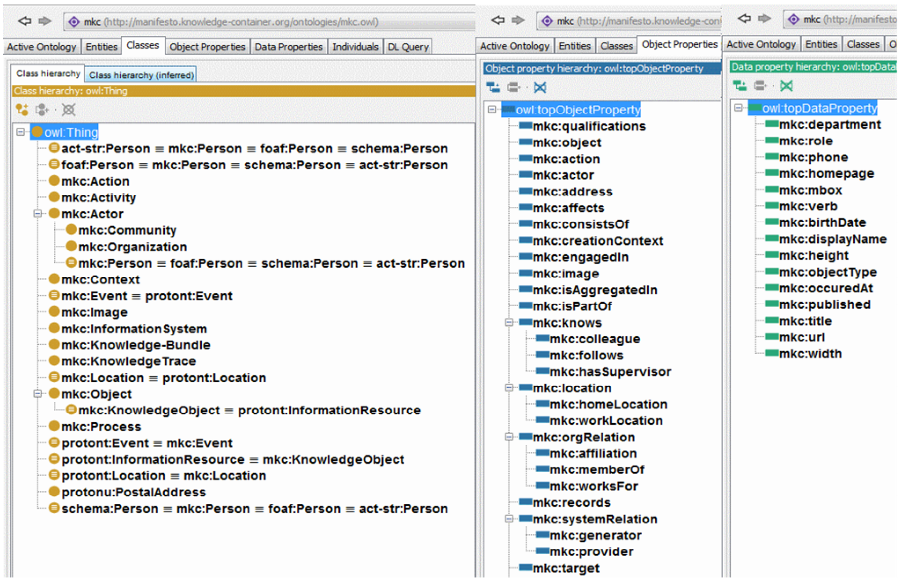

Fig. 1. Screenshot of the OWL ontology in Protégé 4.1

The description of knowledge workers that serve as actor or target of a KT was mainly taken from the person entity of schema.org and activitystrea.ms. FOAF does also have similar attributes for persons and was taken into account as well. All of the mentioned standards are missing a description of a knowledge worker's skills or qualifications. Therefore, we had a look at HR XML, a standard for exchanging information about human resources and used their approach for describing qualifications.

Another challenge was how to cope with the "semantification" of existing XML standards. The verbs and object types of the activitystrea.ms standard for example are defined as strings, which makes sense on an XML level. On the OWL level, it would be better to define them as individuals of the class Action, we have modeled. This issue is not yet solved and should be discussed in a wider audience.

For IT systems we aimed at using BPEL since we are coping mainly with Web-based systems that offer their functionality as a service. However, it is not quite clear which kind of service to reference. Consider a "document created" KT for example. Should we reference a document upload service, a create new document service or a retrieve document service? All three services would describe the same KT being "document created".

4 Validation and Lessons Learned from Implementation

In this section, we first discuss the findings of the general validation results of our anticipatory specification regarding the quality of the proposed standard, before we describe detailed lessons learned (LL) as input for future development and standardization efforts.

Our first proposal of the ontology contains currently 18 classes, 30 object properties, and 15 data properties. Compared to other ontologies this might be still relatively small which is due to the early state of the development. The Attribute Richness (AR=0.83), which is defined as the average number of attributes per class, also indicates that the amount of knowledge about classes is still low in our ontology and needs further enhancement [36]. However, when calculating the relationship richness (IR=0.88), which represents the ratio between the relationships defined in an ontology and the sum of subclasses plus number of relationships [36], our ontology tends to have a high diversity of relations other than class-subclass relations. However, these metrics describe more the nature or characteristic than the efficiency or inefficiency of an ontology and should therefore be carefully interpreted. Additional improvements and enhancements as well as the existence of a large amount of instances will allow more reliable and detailed quality metrics in the future.

As indicated in section 3 we adopted assessment metrics from RMAG [15] and applied them in our technical validation (Table 1).

Table 1. Assessment Perspectives adopted from Reference Model Analysis Grid (RMAG) [15]

Dimension	Validation result
Suitability	Our proposal allowed us to describe the artifacts derived from the scenarios in sufficient detail to be helpful in reconstructing the creation context of information. We built on OWL, XML and other standards.
Conceptual support	Data models [11] and technical implementation (section3.1) have been validated against common frameworks such as the CEN framework [9] or global KM framework [10]: Main elements are supported.
Tools	Currently, there are only generic tools and no effort was taken yet to integrate our standard into existing tools. However, we identified a number of new requirements for tool support such as filtering and versioning (see below).
Development and maintenance of further models and standards	It could be shown that our model [11] is easily adaptable and extensible, e.g., for domain-specific KM applications. The model has been designed and is intended for being extended using XML and OWL.
Business and Management	Our validation has shown that common scenarios (section 3.1) and corresponding processes are covered. However, additional scenarios are needed to be evaluated to show a universal applicability.
Dependence	We only used open and well-adopted standards [11]. Despite that, we did not include big IT vendors nor standardization bodies in our development yet.
Openness	The reuse of existing standards like FOAF, activitystrea.ms and OpenDocument has been key to the development of our proposal. Mappings from our concepts to existing ones provide links.
Expressive power	For the presented scenarios, the context could be represented in our specification. In addition to that, we identified missing concepts and expressiveness in existing standards, like versioning in RDF.
Completeness	Our scenarios (section3.1) could be modelled. Further scenarios have to be analyzed in order to prove a wide applicability.
Technical interoperability	Basic interoperability is granted through the use of OWL and XML, as well as OOXML/OpenDocument. On a semantic layer, interoperability is achieved through mappings of existing standards in the spirit of open linked data.
Understandability	Our specification is still intended for developers and not yet for end users. Human readability of the storage format is somewhat restricted by RDF (see below). Extensive description of the ontology concepts are available.
Coherence and non-redundancy	In the design phase (section 3.1 & 4), we have analyzed the utilized standards and eliminated redundant elements. Elements like displayName, title and URL are shared across classes. All kinds of relations allowed by OWL were used.

This initial validation shows that the proposed standard is suitable for the scenarios analyzed and fulfills most quality requirements for standards. During the assessment of the utility of our artifacts, we generated a set of LL. Each LL focuses on a criterion, which we identified as critical for our approach (e.g., technical interoperability and deployment). In the following we provide more detailed descriptions of these LL. In order to present a glimpse on potential solutions we added short conclusions or (if possible) recommendations to overcome the obstacles taken into account.

Versioning. A challenge that arose from implementation of KT was how to cope with the problem of versioning. Since actors and KOs referenced in KT are typically used multiple times they should be referenced only once in the KT. Otherwise, they would have to be stored redundantly along with display name and other relevant attributes. In addition to that, some of the attributes might change during the life cycle. Therefore, the KT should represent a snapshot of the data that was current at the time when the KT was captured. However, it should be clear to see, that the different versions of the actor or KO are all just versions of the same object and not different objects. We therefore envisioned to store the attributes locally within our RDF/XML structure of the KT and additionally reference the latest version of the actor/KO that should be stored separately in the KC. Unfortunately, as it turned out, this approach is not compliant with RDF rules that either allow an inline description of an RDF triple's object or a reference to an existing resource stored somewhere else in the RDF file or even in a separate file. It was not possible to use both mechanisms at the same time. Also, literature on RDF versioning revealed a lack of research on a solution to do exactly that. On the level of OWL there is the "same individual" construct which can be used for stating, e.g., that "Chancellor Merkel" and "Angela Merkel" are the same, but again, this was not meant to be used for different versions of the same resource. [37] suggests to use RDF collections or containers, specifically the rdf:Alt container to address this issue. In our tests, this was somewhat working, although it led to relatively complex RDF graphs and the W3C RDF validator for example failed to render the respective RDF graph for a small example.
LL 1: RDF does not adequately support version control of different knowledge containers.

Heterogeneous identifiers. Another LL was that if KO as well as persons or other resources are exposed as RESTful Web Services, the URLs can directly be reused as URIs for identifying the RDF metadata. Other ID like UUID or NTLM usernames for users have to be stored in a separate attribute or prefixed with a URN scheme so that they conform to the URI standard.
LL 2: Identifiers other than URI have to be stored in a separate variable or aligned to the URN scheme.

Human readability of KC. The XML structure for RDF metadata differs from the usual structure one would use for an XML document that stores the same information. RDF requires to build the hierarchy using references with rdf:resource to rdf:about, whereas XML utilizes nested XML tags. Although this makes no big difference when an application parses the file, it significantly hinders human readability, which was an important secondary objective for the original design of XML [38]. Since RDF was designed for machine readability, this is not important for RDF despite its use of XML as a serialization format.

LL 3: Human readability of the knowledge container may be restricted under the use of RDF.

Integration of existing standards. As the number of standards available for reuse for our efforts is abundant, we first tried to adopt the respective elements defined and import the namespaces directly. However, this lead to a large number of XML namespaces which complicated the readability. Since most of the standards are specified as XML schema instead of RDF schema or OWL, it also made no sense to use them as namespace for RDF elements, although an RDF parser would accept that. Therefore, we decided to define an own vocabulary that uses the same names for elements and provides mappings either on OWL level or as an XSL transformation.

LL 4: The integration of the various standards in a unified KM standard is challenging as their underlying schemes (XML Schema, RDF Schema, OWL, HTMLmicrodata) differ significantly.

Filter mechanisms. When we mapped the KT of the scenarios to our standard, we soon realized that the knowledge worker who accesses a KO might need a personalized subset of KT. Our KM standard and tools supporting it should therefore facilitate filter mechanisms, which offer different views on the KC. A good starting point to personalize the KC may be the creation of a user profile that is based on the activity stream of the user [39]. This may be used to filter KT that are not corresponding to interests of the user. However, prerequisites to select KT properly are detailed metadata descriptions of the knowledge worker including the social network and the KA at the time when the KT was recorded. Current standards only focus on different aspects, such as describing personal identities (schema.org/Person), social networks (XFN) or streams of events in knowledge activities (activitystrea.ms) and thus do not provide sufficient information for complex filter processes when applied independently.

LL 5: Filter mechanisms are required to address the specific needs of the knowledge worker.

Aggregation. Another possibility to overcome the information overload caused by the abundance of KT in a KC may lie in aggregating several KT to one compound KT. This can be depicted in the new product development scenario. On the one hand, we may aggregate KT of several persons, which perform a specific type of action to a KO, to one action-oriented KT. For example, “Kurt, Elaine, and Justin commented the technical solution paper”. On the other hand, we may aggregate KT belonging to a specific person, who performs several actions on the KO, to one person-oriented KT. For example, “Kurt creates, adds an image to, and comments the technical solution paper”. A controlled vocabulary of possible actions and objects as it is provided in the activitystrea.ms standard may support the process of building collections of KT. Research on enterprise activity streams has already shown the importance of social navigation and aggregation features [40]. Users that are interested in a certain KO mostly refer to persons that are associated to the KO or actions that were performed on the KO.

LL 6: Different levels of KT aggregations are needed to overcome information overload.

Assigning KT to KO. In almost all scenarios it turned out to be difficult to establish a relationship between a KT and its corresponding KO. For example, “Kurt schedules a videoconference with Elaine, Terry and Justin. During the meeting he creates a meeting minutes that summarize the outcomes”. The KT “create document” or “add section in document” could be automatically assigned to the corresponding KO. The former KT “schedule a videoconference”, however, and its related voice and chat protocol may currently not be easily associated to the minutes document. One possibility to overcome this problem is manually assigning video and chat protocol to the minutes document. Nevertheless, the aim of the proposed standard is to achieve a metadata tagging, which is carried out as automatically as possible.

LL 7: Not all KT may be automatically assigned to its corresponding KO.

KM appropriateness. The action verbs specified in the activitystrea.ms standard are tailored to the domain of social network services and do not sufficiently describe the actions which are usually triggered in KA. For example, the verb "checkin" is used for places and events and not for applications, such as "check in document in a system". For the opposite "check-out" event there is no equivalent in activitystrea.ms. The closest semantic construct would be person A (actor) assigned (verb) document (object) to person A (target, person herself). There is also no verb for generating a new version of a document, just update or add. Finally there are like and dislike as verbs to express a personal preference in Facebook style, but no equivalent for expressing 5-star ratings which are common in other systems. The same could be argued for object types. Although the standard defines about 30 different types, there is no differentiation between different kinds of documents like presentation, business letter, project proposal, calculation and so on, only file and the option to associate a mime type.

LL 8: Current Standards do not adequately focus on specific aspects of KM.

Handling of nested KO. Within the research scenario, all five experts share and create a variety of sub-KOs, such as literature, figures, notes or sketches. These sub-KOs have their own bundles of KT that should be kept when included in a KO of a higher aggregation level. When referencing KOs to other KOs by using URL links, this might not be a problem, since all KT can be derived easily. However, if a KO is manually linked by using copy and paste features of current software systems, the maintenance of all corresponding KT cannot be assured.

LL 9: Future software applications have to consider the KT of referenced, linked and nested KOs sufficiently.

Modeling of RDF containers in OWL. During modeling of the ontology and creating corresponding RDF/XML instances, we found that although in RDF it is quite common to use rdf:Bag or parseType=Collection in order to specify a collection of multiple elements, in OWL there is no good way to express such a container structure. If you try to assign rdf:Bag as the range of an object property, it is marked as error by Protégé 4.1, the tool we used.

LL 10: OWL enforces flat structures, whereas XML and RDF allow encapsulating similar elements in a container.

Our key findings thus show the suitability as well as challenges for the standardization process. While feasible on a general level, several deficits still need to be addressed. This should take place in a standardization process towards a broad consensus. This process needs to take our findings into account but also needs to consider stakeholders' (and their systems') requirements as a next step.

5 Conclusion

This paper motivates for joining our initiative to promote standardization in the field of KM in general and in social environments in particular. The manifold developments in social computing have caused a new wave of efforts in organizations to support ad-hoc, grassroots and self-organized activities by knowledge workers to build or appropriate tools that support them in their activities at the workplace. At the moment knowledge integration is hampered by contributions in diverse social environments that fail to be meaningfully integrated. Our initiative for standardization aims to bridge these kinds of gap not only between social environments within organizations, but also between organizational social environments and social environments from business partners as well as global social environments that span the boundaries and are targeted at connecting individuals throughout the world. The resulting specification reuses existing standards and is technically implemented as an ontology (section 3.2) striving to set forth efforts in creating a standard in the domain of KM. The paper reported on selected assessment perspectives taken for initial validation which were adopted from the Reference Model Analysis Grid [15]. In addition – as validating an anticipatory standard is quite complex (section 3) – the technical implementation of three scenarios abstracted from real-world organizations elicited challenges that were discussed as lessons learned (section 4).

The main findings concern technical issues of existing standards that need to be overcome with a unifying specification as well as user needs that up to now have been unaddressed by standards. The lessons learned (LL) identified pave the way for future research in this domain. Next steps should be to implement prototypes to incorporate our standard into existing applications like OpenOffice or MS Office on a technical level, as well as hand the specification over to an international standardization body in order to stimulate further discussion.

6 Acknowledgements

This research has been co-funded by the European Commission under the Information and Communication Technologies theme of the 7th Framework Programme, Integrating Project ARISTOTELE (contract no. FP7-257886) and LAYERS (contract no. 318209).

References

1. Schuler, D.: Social Computing. Communications of the ACM, 37(1), 28–29 (1994)
2. Wang, F.-Y., Zeng, D., Carley, K.M., Mao, W.: Social Computing: From Social Informatics to Social Intelligence. IEEE Intelligent Systems, 22(2), 79–83 (2007)
3. Coates, T.: An addendum to a definition of Social Software. URL: http://www.plasticbag.org/archives/2005/01/an_addendum_to_a_definition_of_social_software/

4. Klamma, R., Chatti, M.A., Duval, E., Hummel, H., Hvannberg, E.T., Kravcik, M., Law, E., Naeve, A., Scott, P.: Social Software for Life-long Learning. Educational Technology & Society, 10(3), 72–83 (2007)
5. Jackson, P.: Capturing, structuring and maintaining knowledge: a social software approach. Industrial Management & Data Systems, 110(6), 908–929 (2010)
6. deVries, H. J.: Standardization: A Business Approach to the Role of National Standardization Organizations, Kluwer Academic Publishers, Boston (2003)
7. Weitzel, T., Beimborn, D., König, W.: A Unified Model of Standard Diffusion: The Impact of Standardization Cost, Network Effects, and Network Topology. MIS Quarterly 30 (Special Issue), 489-514 (2006)
8. Tyrväinen, P., Warsta, J., Seppänen, V.: Evolution of Secondary Software Businesses: Understanding Industry Dynamics. In León, G., Bernardos, A., Casar, J., Kautz, K., DeGross, J. (eds.): IFIP International Federation for Information Processing. Vol 287, Open IT-based innovation: moving towards cooperative IT transfer and knowledge diffusion. Springer, Boston, 281-401 (2008)
9. CEN CWA 14924 European Guide to good Practice in Knowledge Management - Part 1 to 5, Brussels (2004)
10. Pawlowski, J.M. Bick, M.: The Global Knowledge Management Framework: Towards a Theory for Knowledge Management in Globally Distributed Settings. Electronic Journal of Knowledge Management, 10(1), 92–108 (2012)
11. Thalmann, S., Peinl, R., Hetmank, L., Kruse, P., Seeber, I., Maier, R., Pawlowski, J.M., Bick, M.: Manifesto for an Ontology-based Standard on Knowledge Exchange in Social Knowledge Management Environments. 12th International Conference on Knowledge Management and Knowledge Technologies (iKNOW), Graz (2012)
12. Baskin E., Krechmer, K., Sherif, M.: The Six Dimensions of Standards: Contribution Towards a Theory of Standardization", in: Lefebvre, L.A., Mason, R.M., Khalil, T. (eds.): Management of Technology, Sustainable Development and Eco-Efficiency, Amsterdam, Elsevier (1998)
13. Bick, M., Hetmank, L., Kruse, P., Maier, R., Pawlowski, J.M., Peinl, R., Schoop, E., Seeber, I., Thalmann, S.: Manifesto for a Standard on Meaningful Representations of Knowledge in Social Knowledge Management Environments. Multikonferenz Wirtschaftsinformatik (MKWI), Braunschweig (2012)
14. Peinl, R., Thalmann, S., Hetmank, L., Kruse, P., Seeber, I., Pawlowski, J.M., Bick, M., Maier, R., Schoop, E.: Manifesto for a Standard on Knowledge Exchange in Social Knowledge Management Environments. 13th European Conference on Knowledge Management (ECKM), Cartagena (2012)
15. Pawlowski, J.M., Kozlov, D.: Analysis and Validation of Learning Technology Models, Standards and Specifications: The Reference Model Analysis Grid (RMAG), International Journal of IT Standards and Standardization Research, 8 (2), 1-19 (2010)
16. Ma, M., Agarwal, R.: Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities. Information Systems Research, 18(1), 42–67 (2007)
17. Wiig, K.M.: What future Knowledge Management users may expect. Journal of Knowledge Management, 3(2), 155–165 (1999)
18. Shen, J., Fitzhenry, E., Dietterich, T.G.: Discovering Frequent Work Procedures from Resource Connections. In: Proceedings of the 14th international conference on Intelligent user interfaces, pp. 277–286. ACM Press, New York (2009)
19. Rath, A.S., Devaurs, D., Lindstaedt, S.N.: UICO: An Ontology-Based User Interaction Context Model for Automatic Task Detection on the Computer Desktop. In: Proceedings

of the 1st Workshop on Context, Information and Ontologies (CIAO 2009), Article No. 8, ACM, Heraklion, Greece (2009)

20. Riss, U.: Pattern-Based Task Management as Means of Organizational Knowledge Maturing. International Journal of Knowledge-Based Organizations, 1(1), 20–41 (2011)
21. Maier, R.: Knowledge Management Systems: Information and Communication Technologies for Knowledge Management. Springer, Berlin (2007)
22. Cano, A., Tucker, S., Ciravegna, F.: Capturing entity-based semantics emerging from personal awareness streams. In: Proceedings of the 1st Workshop on Making Sense of Microposts (MSM 2011). pp. 33–44 (2011)
23. Freyne, J., Berkovsky, S., Daly, E.M., Geyer, W.: Social networking feeds: recommending items of interest. In: Proceedings of the fourth ACM conference on Recommender systems. pp. 277-280 (2010)
24. Kelloway, E.K., Barling, J.: Knowledge Work As Organizational Behavior. International Journal of Management Reviews, (2)3, 287–304 (2000)
25. Maier, R., Hädrich, T., Peinl, R.: Enterprise knowledge infrastructures. (2nd ed.) Springer, Berlin (2009)
26. Majchrzak, A., Malhotra, R.J.: Perceived individual collaboration know-how development through information technology-enabled contextualization: Evidence from distributed teams. Information Systems Research, 16(1), 9–27, (2005)
27. Maaninen-Olsson, E., Müllern, T.: A contextual understanding of projects—The importance of space and time. Scandinavian Journal of Management, 25(3), 327–339, (2009)
28. Hanakawa, N., Matsumoto, K.-I., Torii, K.: A Knowledge-Based Software Process Simulation Model. Annals of Software Engineering, 14(1-4), 383–406 (2002)
29. Hsu, T.-Y, Ke, H.-R., and Yorkang, W.-P.: Unified knowledge-based content management for digital archives in museums. The Electronic Library, 24(1), 38–50 (2006)
30. Grebner, O., Ong, E., Riss, U., Brunzel, M., Bernardi, A., Roth-Berghofer, T.: Task Management Model - Deliverable D3.1. Artificial Intelligence. nepomuk.semanticdesktop.org/nepomuk/D3-1.html (2006)
31. Tramp, S., Ermilov, T., Frischmuth, P., Auer, S.: Architecture of a Distributed Semantic Social Network. In: Federated Social Web Europe 2011, Berlin, Germany (2011)
32. Hevner, A., March, S., Park, S.T., Ram, S.: Design Science in Information Systems Research. MIS Quarterly, 28, 75–105 (2004)
33. Bria, A., Gessler, F., Queseth, O., Stridh, R., Unbehaun, M., Wu, J., Zander, J., Flament, M.: 4th Generation Wireless Infrastructures – Scenarios and Research Challenges. IEEE Personal Communications 8, 25-31 (2001)
34. de Jong, M., Lentz, L.: Scenario Evaluation of Municipal Web Sites: Development and use of an Expert-focused Evaluation Tool. Government Information Quarterly, 23, 191-206 (2005)
35. Chen, R., Sharman, R., Chakravarti, N., Rao, H.R., Upadhyaya, S.J.: Emergency Response Information System Interoperability: Development of Chemical Incident Response Data Model. Journal of the Association for Information Systems, 9, Article 7 (2008) Available at: http://aisel.aisnet.org/jais/vol9/iss3/7
36. Tartir, S., Arpinar, I.B., Moore, M., Sheth, A.P., Aleman-Meza, B.: OntoQA: Metric-based ontology quality analysis. IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources (2005)
37. Lohfink, A., McPhee, D.: Resource-Level Versioning in Administrative Geography RDF Data. In: Semantic Web – Interoperability, Usability, Applicability. http://www.semantic-web-journal.net/sites/default/files/swj214.pdf

38. Lawrence, R.: The space efficiency of XML. Information and Software Technology, 46(11), 753-759 (2004)
39. Guy, I., Ronen, I., Raviv, A.: Personalized activity streams: sifting through the "river of news". Proceedings of the fifth ACM conference on Recommender systems. pp. 181–188. ACM Press, New York (2011)
40. Shami, N.S., Muller, M., Millen, D.: Browse and discover: social file sharing in the enterprise. Proceedings of the ACM 2011 Conference on Computer supported cooperative work (CSCW 2011). pp. 295–304. ACM Press, New York (2011)