ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTING 172

Lorna Uden
Francisco Herrera
Javier Bajo Pérez
Juan Manuel Corchado Rodríguez (Eds.)

7th International Conference on Knowledge Management in Organizations: Service and Cloud Computing

Springer

Advances in Intelligent Systems and Computing 172

For further volumes:
http://www.springer.com/series/11156

Lorna Uden, Francisco Herrera, Javier Bajo Pérez,
and Juan Manuel Corchado Rodríguez (Eds.)

7th International Conference on Knowledge Management in Organizations: Service and Cloud Computing

Springer

Editors
Lorna Uden
FCET
Staffordshire University
The Octagon, Beaconside
UK

Javier Bajo Pérez
Faculty of Computer Science
Pontificial University of Salamanca
Salamanca
Spain

Francisco Herrera
Department Computer Science and
Artificial Intelligence
University of Granada
Granada
Spain

Juan Manuel Corchado Rodríguez
Department of Computing Science
and Control
Faculty of Science
University of Salamanca
Salamanca
Spain

ISSN 2194-5357
e-ISSN 2194-5365
ISBN 978-3-642-30866-6
e-ISBN 978-3-642-30867-3
DOI 10.1007/978-3-642-30867-3
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012939635

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Knowledge is increasingly recognised as the most important resource in organisations and a key differentiating factor in business today. It is increasingly being acknowledged that Knowledge Management (KM) can bring about the much needed innovation and improved business performance in organisations. The service sector now dominates the economies of the developed world. Service innovation is fast becoming the key driver of socio-economic, academic and commercial research attention. Knowledge Management plays a crucial role in the development of sustainable competitive advantage through innovation in services. There is tremendous opportunity to realise business value from service innovation by using the knowledge about services to develop and deliver new information services and business services.

Although there are several perspectives on KM, they all share the same core components, namely: People, Processes and Technology. Organisations of all sizes across nearly every industry are seeking new ways to address their knowledge management requirements. Cloud computing offers many solutions to the problems facing KM implementation. Cloud computing is an emerging technology that can provide users with all kinds of scalable services, such as channels, tools, applications, social support for users' personal knowledge amplification, personal knowledge use/reuse, and personal knowledge sharing.

The seventh International Conference on Knowledge Management in Organizations (KMO) offers researchers and developers from industry and the academic world to report on the latest scientific and technical advances on knowledge management in organisations. It provides an international forum for authors to present and discuss research focused on the role of knowledge management for innovative services in industries, to shed light on recent advances in cloud computing for KM as well as to identify future directions for researching the role of knowledge management in service innovation and how cloud computing can be used to address many of the issues currently facing KM in academia and industrial sectors. This conference provides papers that offer provocative, insightful, and novel ways of developing innovative systems through a better understanding of the role that knowledge management plays.

The KMO 2012 proceedings consist of 53 papers covering different aspects of knowledge management and service. Papers came from many different countries including Australia, Austria, Brazil, China, Chile, Colombia, Denmark, Finland, France, Gambia, India, Japan, Jordan, Netherlands, Malaysia, Malta, Mexico, Netherlands, New Zeland, Saudi Arabia, Spain, Slovakia, Slovenia, Taiwan, Turkey, United States of America and United Kingdom. We would like to thank our program committee, reviewers and authors for their contributions. Without their efforts, there would be no conference and proceedings.

Salamanca
July 2012

Lorna Uden
Francisco Herrera
Javier Bajo Pérez
Juan Manuel Corchado Rodríguez

Organization

Conference Chair

Dr. Lorna Uden	Staffordshire University, UK

Program Chair

Dr. Francisco Herrera	University of Granada, Spain

Program Committee

Dr. Senen Barro	University of Santiago de Compostela, Spain
Dr. Hilary Berger	UWIC. England. UK
Dr. Pere Botella	Polytechnic University of Catalonia – BarcelonaTech, Spain
Dr. Vicente Botti	Polytechnic University of Valencia, Spain
Dr. Senoo Dai	Tokyo Institute of Technology, Japan
Dr. Flavius Frasincar	Erasmus University Rotterdam, The Netherlands
Dr. Wu He	Old Dominion University, USA
Dr. Marjan Hericko	University of Maribo, Slovenia
Dr. George Karabatis	University of Maryland, Baltimore County, USA
Dr. Dario Liberona	University of Santiago, Chile
Dr. Remy Magnier-Watanabe	University of Tsukuba, Tokyo, Japan
Dr. Victor Hugo Medina Garcia	Universidad Distrital Francisco José de Caldas, Colombia
Dr. Marja Naaranoja	Vaasa University of Applied Sciences, Finland
Dr. Paulo Novais	University of Minho, Portugal
Dr. Takao Terano	Tokyo Institute of Technology, Japan
Dr. Derrick Ting	National University of Kaohsiung, Taiwan
Dr. Lorna Uden	Staffordshire University, UK
Dr. Michael Vallance	Future University Hakodate, Japan

Dr. William Wang	Auckland University of Technology, New Zealand
Dr. Leon S.L. Wang	National University of Kaohsiung, Taiwan
Dr. Guandong Xu	Victoria University, Australia
Dr. Ales Zivkovic	University of Maribo, Slovenia

Organization Committe

Dr. Juan M. Corchado (Chair)	University of Salamanca, Spain
Dr. Javier Bajo (Vice-chair)	Pontifical University of Salamanca, Spain
Dr. Sara Rodríguez	University of Salamanca, Spain
Dr. Francisco de Paz Santana	University of Salamanca, Spain
Dr. Emilio S. Corchado	University of Salamanca, Spain
Dr. Belén Pérez	University of Salamanca, Spain
Fernando De la Prieta	University of Salamanca, Spain
Carolina Zato	University of Salamanca, Spain
Antonio J. Sanchez	University of Salamanca, Spain

Contents

Section 1: Innovation in Knowlege Management

Section 2: Knowledge Management in Business

Section 5: Knowledge Management and Service Science

Section 6: Technology Applied to Knowledge Management

Evaluation of a Self-adapting Method for Resource Classification in Folksonomies

José Javier Astrain, Alberto Córdoba, Francisco Echarte, and Jesús Villadangos

Dept. Ingeniería Matemática e Informática, Universidad Pública de Navarra, Campus de Arrosadía, 31006 Pamplona, Spain
josej.astrain@unavarra.es, alberto.cordoba@unavarra.es, patxi@eslomas.com, jesusv@unavarra.es

Abstract. Nowadays, folksonomies are currently the simplest way to classify information in Web 2.0. However, such folksonomies increase continuously their amount of information without any centralized control, complicating the knowledge representation. We analyse a method to group resources of collaborative-social tagging systems in semantic categories. It is able to automatically create the classification categories to represent the current knowledge and to self-adapt to the changes of the folksonomies, classifying the resources under categories and creating/deleting them. As opposed to current proposals that require the re-evaluation of the whole folksonomy to maintain updated the categories, our method is an incremental aggregation technique which guarantees its adaptation to highly dynamic systems without requiring a full reassessment of the folksonomy.

1 Introduction

Folksonomies are nowadays a widely used system of classifying information for knowledge representation [10]. Tags made by users provide semantic information that can be used for knowledge management. As users annotate the same resources, frequently using the same tags, their semantic representations for both tags and annotated resources emerge [12, 13, 15]. Folksonomies are based on the interaction of multiple users to jointly create a "collective intelligence" that defines the semantics of the information. Although users follow an easy and simple mechanism to classify resources, knowledge representation becomes more difficult as the volume of information increases [3]. Folksonomies are difficult to be studied due to their three-dimensional structure (hypergraph) [2, 9]. Then, different two-dimensional contexts of this information are often considered [2] (tag-user, tag-resource and tag-tag).

Folksonomies are highly dynamic systems, so any proposal should be scalable and able to adapt to its evolution. In [8], a generalization of the methods used to obtain two-dimensional projections dividing them into non-incremental aggregation methods and incremental aggregation methods is proposed. The first one includes solutions similar to those proposed in [2, 9] where the incorporation of new information to the folksonomy involves the complete recalculation of the similarity matrices. The second one includes solutions in which new annotations introduced in

L. Uden et al. (Eds.): 7th International Conference on KMO, AISC 172, pp. 1–12.
springerlink.com

the folksonomy do not involve a repetition of all the calculations. Faced with these tag-centric based proposals, resource-centric approaches have been poorly studied with the aim of structuring the folksonomy resources. In this paper, we consider the improvement of the folksonomy knowledge representation by creating semantic categories, also called concepts, that group the folksonomy resources. We try to obtain the relationship between kolksonomies and ontologies obtaining the semantics from both tag and resources.

Some works in literature [1, 14] attempt to classify the resources of a folksonomy into concepts to offer improvements in user navigation. In [1] the resources of a folksonomy are classified under a set of classification concepts. These classification concepts are previously obtained through a manually search through a repository of ontologies. Once obtained the classification concepts, an algorithm classifies the tags of the folksonomy under the concepts obtained combining the co-occurrences of the tags and the results obtained after submitting certain search patterns to Google. Folksonomy resources are classified into concepts according to the tags they have been assigned. Furthermore, the authors do not propose an implementation or prototype of their proposal, so it is not possible to assess to what extent the classification aided navigation. This non-incremental method depends on the intervention of a user which defines the terms for the classification of ontologies in which resources are classified, and the use of external information sources. Authors also fail to take into account the evolution of the folksonomy and how to adapt their proposal to the incorporation of new tags and resources.

An optimization algorithm for the classification of resources under a set of concepts, using a set of predefined concepts and a set of previously classified resources, is used in [14]. In this regard, the goal of this algorithm is similar to the method described in this work, since it directly classifies the resources under concepts. However, the algorithm has some drawbacks such as: a) the categories are fixed, and their evolution are not considered; b) it requires full reassessment of the algorithm each time the folksonomy evolves since it is a non-incremental method; and c) it does not describe how resources are sorted within each concept.

This paper introduces a simple method for the automatic classification of the resources of a folksonomy into semantic concepts. The main goal is to improve the knowledge management in folksonomies, keeping the annotation method as simple as usual. Folksonomies are highly dynamic systems where new tags and resources are created continuously, so the method builds and adapts these concepts automatically to the folksonomy's evolution. Concepts can appear or disappear by grouping new resources or disaggregating existing ones and resources would be automatically assigned to those concepts. Furthermore, the method has a component-based open architecture which allows its application to folksonomies with different characteristics. It uses a reduced set of the folksonomy's tags to represent both the semantics of the resources and concepts because it requires a high classification efficiency to allow its application to real folksonomies (where the number of annotations grows very fast), and assigns automatically an appropriated name to those concepts.

The rest of the paper is organized as follows: Section 2 describes the method; Section 3 deals with the method evaluation using a real folksonomy; Acknowledgements, Conclusions and References end the paper.

2 Method Description

The method, introduced in [6], follows a component based open architecture, which allows its application to folksonomies with different characteristics. It requires a high classification efficiency to allow its application to real folksonomies, where the number of annotations grows very fast.

Given a folksonomy, the method initially creates a set of concepts where resources are grouped, and it assigns a name to each concept according to the semantic information provided by the resources grouped in each concept and their annotations. Once created the concepts, each new annotation of the folksonomy is processed updating the semantic information and adapting the concepts when necessary. The method starts with the creation of the set of representative tags (S_{rt}) and the vectorial representation of the resources of the folksonomy. The component *Representations* is in charge of these tasks. It may use different criteria to create S_{rt} like the tags more frequently used, the tags used by more users and even more. Then, each resource is assigned to subsets $R_{converged}$ or $R_{pending}$ in terms of whether they have converged or not. In order to obtain those tags that best describe the semantic of a resource, in [11], it is showed that tagging distributions of heavily tagged resources tend to stabilize into power law distributions. The component *Convergence* may use many criteria like the total amount of annotations, or the number of annotations associated to the S_{rt} set to assign the resources to $R_{converged}$.

The component *Clustering* clusters the resources of the folksonmy belonging to $R_{converged}$ in a set of concepts, generating the set of semantic concepts on which resources of the folksonomy are grouped (C) and the set of pairs (r, c) where $r \in R$ and $c \in C$, representing that resource r is grouped into the concepts c (Z). The initial clustering of the resources may follow different criteria: applying clustering techniques to the resources of the folksonomy, creating manually the classification concepts and selecting a relevant resource as seed of the classifier, or adapting some tag clustering algorithms like T-Know [1] in order to classify resources, instead of tags, under the concepts that have been previously selected. The component *MergingSplitting* analyses the concepts provided in order to evaluate the convenience of merging or splitting any of them, updating C and Z sets. It merges those concepts whose similarity values are greater than 0.75, using the cosine measure. The component *Classifier* is in charge of grouping the resources under those concepts with which they have high semantic similarity by comparing all the resources belonging to the $R_{converged}$ set with the concepts in which they are grouped in Z. The component assigns the resource to the $R_{classified}$ set according to the similarity measures between each resource and its group or keeps it on $R_{converged}$ and removes it from Z. The component *Representations* creates the vector representations for the C concepts using the Z set, and the component *Naming* assigns meaningful names to these concepts according to some criteria like the tags with higher weights.

```
1.  Representations::createSrt()
2.  Representations::createVectors()

3.  forall r ∈ R do
4.      if Convergence::hasConverged(r) then
5.          assign r to R_converged
6.          else assign r to R_pending
7.      endif
8.  endforall

9.  Clustering::create(R_converged)
10. MergingSplitting::process(C,Z)

11. forall r ∈ R_converged do
12.     if Classifier::isCorrectlyClassified(Z,r) then
13.         assign r to R_classified
14.         else drop r from Z
15.     endif
16. endforall

17. Representations::createConceptVectors(C,Z)
18. Naming::process(C,Z)

19. while true do

20.     tagging = wait(FolksonomyEvolution)

21.     if not Representations::inSrt(t) then
22.         continue
23.     endif

24.     Representations::updateVectors(Tagging)

25.     if tagging.action = create then
26.         if r ∈ R_pending
                and Convergence::hasConverged(r) then
27.             assign r to R_converged
28.         endif
29.         if r ∈ R_converged then
30.             Classifier::classify(Z,r)
31.             Representations::updateConceptVector(Z,r)
32.         endif
33.     endif

34.     if RecalculationCondition::check() then
35.         Representations::updateSrt()
36.         Clustering::update(C,Z)
37.         MergingSplitting::process(C,Z)
38.         Representations::updateVectors(C,Z)
39.         Naming::process(C,Z)
40.     endif

41. endwhile
```

Fig. 1 Method algorithm

At this point, the method has built $R_{pending}$, $R_{converged}$, $R_{classified}$, C and Z sets from the folksonomy, so it provides a set of concepts that group folksonomy resources based on their semantics. Once created these concepts, the method self-adapts to the evolution of the folksonomy taking into account the new annotations made by users. The lines 19 to 41 of the pseudocode described in Fig. 1 are responsible of processing these new annotations. The method waits for a change on the folksonomy when creating or removing an annotation. This change is represented by the method as a new *Tagging* element, which contains the annotation information (user, resource and tag), and if it has been created or deleted. If the tag used in the *Tagging* does not belong to the representative set of tags (S_{rt}), *Tagging* is ignored and it expects the reception of new annotations. If this tag belongs to the S_{rt} set, the component $Representations$ updates the vectorial representation of the resource. If the resource belongs to the $R_{classified}$ set the component also updates the vectorial representation of the concept in which the resource is grouped. If the *Tagging* is of type *create*, the method checks whether the resource has converged or not, and the possibility of grouping it under any existing concept. If the resource belongs to the $R_{pending}$ set, the component *ConvergenceCriterion* checks if the resource has converged after receiving the new annotation. If so, or if the resource already previously belonged to $R_{converged}$, the component *Classifier* provides the most appropriate concept for this resource. The component compares the semantic of the resource with that of each concept in C. Based on this similarity, the component assigns the resource to the $R_{classified}$ set and creates a new entry in Z, or lets the resource continue to be assigned to $R_{converged}$. If the resource is assigned to a concept, *Representations* component updates the vectorial representation of the concept with the resource information.

The method groups the resources in concepts, so that once a resource is classified it will never return to the set $R_{converged}$. Therefore, when a *Tagging* of type *delete* is received, the method does not checks if the resource has converged or whether it must continue to exist under the current concept, the method only updates the corresponding vectorial representation. In addition to gathering new converged resources into existing concepts, the method considers the information received from the new *Taggings*, updating the S_{rt} set and the existing concepts. Thus, S_{rt} and C sets may adapt to the folksonomy's evolution, performing their adaptation for example to new users' interests. Since a unique *Tagging* does not use to significantly affect the S_{rt} set or the concepts set, and this update can be quite expensive computationally, the method uses the component *RecalculationCondition* to determine when to update both concepts and S_{rt}. The recalculation may be performed considering many criteria, for example, after a certain number of processed *Taggings*, after a given time period, or whenever a resource or a set of resources are classified. When the component determines the convenience of performing the recalculation, in a first step the S_{rt} set is updated taking into account its establishment criteria using component *Representations*. It then uses the *Clustering* component to update the existing concepts (C) and the resources grouped in them (Z). The component *MergingSplitting* reviews these concepts creating, splitting them when necessary. Once obtained the elements C and Z, *Representations* updates the concept representation vectors, and *Naming* assigns a name to each one of the concepts. Upon the completion of these tasks, the method returns to stand waiting for the arrival of new *Taggings* to the folksonomy for their processing.

3 Method Evaluation

This section is devoted to evaluate experimentally the proposed method using data retrieved from Del.icio.us. The method creates automatically a set of concepts from an existing folksonomy and groups under these concepts the resources of the folksonomy, according to their semantics. As the folksonomy receives new annotations, the method groups new resources, takes into account new relevant tags, and adapts the existing concepts.

With the aid of a page scraper we have collected a set of 15,201 resources, with 44,437,191 annotations, 1,293,351 users and 709,657 tags from Del.icio.us (15th-30th September, 2009)[1]. Those annotations, depicted in Table 1, concern a period time from January 2007 to September 2009. We use annotations prior to 2009 to simulate an initial state of the folksonomy in order to create the initial concepts. The rest of annotations correspond to January to September 2009 and they are used to simulate the folksonomy evolution by means of *Taggings* elements of type *create*. Table 2 summarizes the information concerning the initial folksonomy (t_0) and its state after including the *Tagging* elements concerning the period Jan.-Sept. (t_1 to t_9).

[1] http://www.eslomas.com/publicaciones/KMO2012/

Table 1 Annotation distribution

Year	Annotations	Year	Annotations	Year	Annotations
1998	2	2002	299	2006	3,140,591
1999	3	2003	628	2007	7,237,129
2000	7	2004	52,345	2008	13,753,922
2001	12	2005	719,216	2009	19,533,037

Table 2 Users, tags and resources in each subset of the experiment

	Increment of the number of elements				Aggregated values			
t_i	annotations	users	resources	tags	annotations	users	resources	tags
t_0	24,904,154	972,695	12,117	489,125	24,904,154	972,695	12,117	489,125
t_1	1,704,682	35,480	342	23,581	26,608,836	1,008,175	12,459	512,706
t_2	1,811,331	37,240	353	23,066	28,420,167	1,045,415	12,812	535,772
t_3	2,179,539	40,672	407	26,101	30,599,706	1,086,087	13,219	561,873
t_4	2,153,461	34,603	336	24,187	32,753,167	1,120,690	13,555	586,060
t_5	2,230,512	33,078	391	24,919	34,983,679	1,153,768	13,946	610,979
t_6	2,304,614	33,959	348	24,460	37,288,293	1,187,727	14,294	635,439
t_7	2,437,317	34,137	345	24,951	39,725,610	1,221,864	14,639	660,390
t_8	2,617,998	36,438	368	26,332	42,343,608	1,258,302	15,007	686,722
t_9	2,093,583	35,049	194	22,935	44,437,191	1,293,351	15,201	709,657

The method is configured for the experimentation using the following components. The comparison between resource and concepts vectors has been performed using the cosine measure. In the following we describe the components used to configure the method. The component *Representations* considers a S_{rt} set built using those tags with at least 1,000 annotations. The component *Convergence* fixes the convergence criterion to 100 annotations [7]. The component *Classifier* uses the method presented in [5] to classify the resources under the most similar concepts. In the evolution task, the classifier applies for a given resource each time it reaches a multiple of 50 annotations. This component has been configured to take into account the two most similar concepts (c_i, c_j) to each resource (r_i), and classify only the resource when the difference between the resource and these concepts is greater than a minimum threshold value of 0.10 ($|sim(r_i, c_i) - sim(r_i, c_j)| \geq 0.10$). As proved in [5], the method is able to classify the resources providing a high precision. The component *Recalculation* recalculates monthly both sets S_{rt} and C (after each t_i $(i : 1..9)$). The component *Clustering* uses a k-means algorithm, determining the k value by the expression $k = \sqrt{\frac{n}{2}}$, being n the number of resources in $R_{converged}$ at the concepts creation and in $R_{classified}$ when recalculating. Thus, the number of concepts can grow as the folksonomy evolves. At the initial concepts creation, the initial centroids are randomly defined since any a-priori knowledge is not considered. When recalculating, the representation vectors of C concepts are used to define the initial centroids, and if $k \geq |C|$, then some resources are randomly selected to define the $k - |C|$ new clusters. The implementation of the algorithm is performed in a distributed way where the calculus of the number of cluster changes at each iteration is performed over different PCs of a cluster using a task queue manager (Gearman). We use the k-means method instead of hierarchical techniques because we want to provide a concept cloud (see Figure 2) with a similar user experience to tag clouds. Component *MergingSplitting* merges using the cosine measure those concepts whose similarity values are greater than 0.75, once the *Clustering* component obtains C and Z sets.

Fig. 2 Concept cloud

And finally, component *Naming* assigns a name to each concept according to the most relevant tags of its vector. When the weight of several tags is greater than the 50% of the weight of the most relevant tag, the name of the concept is obtained through the concatenation of those tags after verifying, by means of the Levenshtein distance, that tags are not syntactic variations of other previous tags. So, a concept in which the two most relevant tags are *php* (weight 127,427) and *programming* (weight 39,743), is named “Php”.

Besides Gearman, Memcached has been used as a cache system in order to reduce the number of accesses to the database to obtain resources, concepts and representation vectors. The employed hardware consists of four commodity PC with Intel Core 2 Duo processors at 2.13 GHz, and 2GB of RAM memory. In order to perform the distributed tasks, 8 processes (2 on each PC) have been executed to perform the processing of the K-means slices, and 24 processes have been executed to process the *Tagging* processing tasks.

Table 3 summarizes the information concerning the evolution of the number of tags, resources and concepts, from t_0 to t_9. Regarding the tags and the S_{rt} set, it shows the evolution of the number of tags of the folksonomy and the number of tags in S_{rt}, and the ratio between them. One can note that the number of tags in S_{rt} increases as the number of annotations in the folksonomy does. The increment in the number of annotations carries with a higher number of tags exceeding the threshold of 1,000 annotations required to be part of S_{rt}. Lets note that: i) the method represents the semantics of the resources with less than 0.40% (1,939/489,125) of the existing tags; and ii) this value decreases slowly as more annotations arrive to the folksonomy (up to 0.38% after processing t_9). This makes the cost of the method, in terms of space and process, significantly lower than the required when considering all the tags of the folksonomy to represent the semantics of the resources and their concepts associated. Table 3 shows that the number of classified resources increases as the number of resources of the folksonomy does. As folksonomy evolves with new annotations, some of the $R_{pending}$ resources converge and pass to $R_{converged}$ set, while other resources receive enough annotations to determine an adequate concept, passing to $R_{classified}$. Regarding concepts, Table 3 also shows the evolution in the k value used by the *Clustering* component, and the number of concepts created at each recalculation. When recalculating, *Clustering* component tries to create a number of k concepts considering the number of $R_{classified}$ resources, however, some of these concepts may be merged by *MergingSplitting* component when their similarity are greater than the defined threshold (0.75). Note that both k and $|C|$ values gradually increases after each recalculation. Although in this experiment the number of concepts increases, the method allows the creation, splitting and merging of concepts, so this number may also decrease. In this experiment the clustering is based on k-means, with k depending on the number of resources, causing the maintenance or increase of the number of concepts. The number of concepts can decrease only if, after clustering, the *MergingSplitting* component finds that there are two or more concepts with a similarity degree greater than 0.75.

Table 3 Adaptation of the method as the folksonomy evolves

	Tags			Resources			Concepts	
t_i	$\|T\|$	$\|S_{rt}\|$	$\frac{\|S_{rt}\|}{\|T\|}$	$R_{pending}$	$R_{converged}$	$R_{classified}$	$\|C\|$	k
t_0	489.125	1.939	0,40%	337	3.098	8.682	75	77
t_1	512.706	2.037	0,40%	289	3.184	8.986	76	78
t_2	535.772	2.124	0,40%	305	3.190	9.317	77	80
t_3	561.873	2.204	0,39%	282	3.226	9.711	78	81
t_4	586.060	2.286	0,39%	269	3.237	10.049	78	82
t_5	610.979	2.362	0,39%	250	3.277	10.419	79	83
t_6	635.439	2.437	0,38%	225	3.300	10.769	80	84
t_7	660.390	2.526	0,38%	206	3.033	11.400	81	85
t_8	686.722	2.616	0,38%	139	2.915	11.953	82	86
t_9	709.657	2.716	0,38%	43	2.917	12.241	83	87

Analysing the creation of concepts and the evolution in the number of resources grouped in each concept, one can note that the creation of certain concepts produces a decrease in the number of resources grouped in other concepts. For example, after processing the *Tagging* set t_9, the new annotations received in the transition from the eighth to ninth iteration lead the concept $Business\&Finance\&Money$ be split into two concepts: $Business$ and $Finance\&Money$. The initial concept ($Business$), which previously brought together 215 resources, now groups 118 resources. The new concept created groups 119 resources, of which 110 (92 %) were previously grouped under $Business\&Finance\&Money$. In most cases, the number of resources classified under each concept grows as the number of annotations of the folksonomy does, but in some cases, this number may decrease. This decrease occurs either when the concept under which resources were classified is split either when the arrival of new annotations allow a better definition of the classification concept. As occurs with $Web2.0\&Social\&Socialnetworking$, in which the number of classified resources decreases from 244 to 230 in the transition from iteration 8 to 9. In this transition, 220 resources remain classified under this concept, 24 of them are reclassified under other existing concepts ($Video$, $Business$, $Twitter$, $Blog$, $Tools\&Web2.0$ and $Generator\&Tools\&Fun$), 6 new resources are classified under this concept and the remaining 4 resources were previously classified under other concepts.

Table 4 Number of resources classified under the eight first concepts

	Concept	Statistics				Iteration									
		Avg	Dev	Max	Min	0	1	2	3	4	5	6	7	8	9
1	Hardware&Electronics&Technology	4.2	64	72	60	60	61	62	63	64	64	66	69	71	72
2	Jobs&Career&Job	85	6.4	95	75	75	77	81	83	84	84	88	90	92	95
3	Books&Literature	171	13.0	192	152	152	156	161	166	169	172	177	182	185	192
4	Video&Web2.0	214	26.9	252	175	175	183	192	201	206	221	225	237	248	252
5	Html&Webdesign&Web	85	9.2	103	75	75	77	80	80	80	81	85	88	98	103
6	Linux&Opensource	116	6.4	125	108	108	109	110	112	115	117	118	122	125	125
7	Art&Design	198	7.0	208	189	189	190	195	193	194	200	204	205	206	208
74	Howto&Diy	81	8.6	93	70	82	70	72	72	75	83	86	88	92	93

Table 5 Evolution of the resources classified

Number of changes	Initial iteration										Total
	t_0	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	
1	320	25	29	17	11	14	11	26	22	0	475
2	79	2	1	3	2	3	0	1	0	0	91
3	2	0	1	0	0	0	0	0	0	0	3
4	1	0	0	0	0	0	0	0	0	0	1

Table 4 shows the number of resources classified under some concepts. In the same way, the creation of the concept *Origami*, which groups 16 resources, occurs after processing the set t_1 and its resources come from *Howto& Diy*, *Art&Design* and the rest are resources classified during the processing of t_1 and therefore were not grouped into any concept after t_0. It has been found therefore that the origin of the resources of the concepts created in the evolution presents a semantic relationship to the concepts created. This indicates that as the folksonomy receives new annotations, the method is able to group all the related resources and to evolve the concepts so as to adapt to user interests: e.g., resources that were initially grouped into *Howto&Diy*, representing pages with instructions on how to do things with the appearance in the folksonomy of more information on origami causes the method to create a specific concept for this topic, bringing together new and existing resources.

Table 5 shows the evolution of the resources classified. The left side of the table shows that 11,671 of the 12,241 resources retain their original classification along the evolution of the folksonomy, 475 of the resources change their concept of classification once, 91 of them change twice, 3 of them make it thrice, and only 1 makes it 4 times. The right side shows the distribution of the number of changes among classification concepts experienced by resources according to the iteration in which they are introduced. Of the set of resources introduced in the initial iteration which change their classification concept along the evolution of the folksonomy, 320 of them make it once, 79 of them change their classification concept twice, and so on. An analysis of the 83 classification concepts involved in this process shows that 75 of them appear in the initial iteration, and the rest appear one for each iteration, except in the fifth iteration, which does not introduce any new concept.

Regarding the performance of the method, the time spent in each stage has been registered. Table 6 shows the time needed to create the initial concepts (lines 1-18 of the method described in Fig. 1).

Table 6 Creation costs for the initial concepts

Action	Time (minutes)
S_{rt} creation	1.10
Vectorial representation creation	22.83
Convergence verification	0.08
Clustering (k=76)	192.21
MergingSplitting	0.98
Assignation to $R_{classified}$	61.62
Creation of concept' vectors	0.53
Naming	0.01

Table 7 Time spent in evolution

		Time (minutes)	
t_i	Number of taggings	Tag. Processing	Recalculation
t_1	1,704,682	28.23	23.47
t_2	1,811,331	31.24	24.35
t_3	2,179,539	41.98	48.91
t_4	2,153,461	43.12	29.18
t_5	2,230,512	40.36	56.45
t_6	2,304,614	41.12	27.39
t_7	2,437,317	41.84	81.34
t_8	2,617,998	32.65	27.26
t_9	2,093,583	37.34	63.12

Table 7 shows the time spent processing the *Tagging* elements of t_1 to t_9 sets (lines 19 to 41 in Fig. 1). Regarding the *Tagging* elements processing, the method has processed them with an average throughput of 965.74 *Tagging* elements per second, which represents an average of 40.24 *Tagging* elements per second and tagging processing worker (24 workers, threads annotating concurrently). These results show the applicability of the method in real systems, processing the new annotations to adapt the classification concepts. Regarding the time spent in the recalculation, it is quite variable, depending on the number of iterations of k-means algorithm. However, the time spent by the *Clustering* component when recalculating is much lower than the time spent at the first clustering (t_0), because the representation vectors of C are used as initial centroids, while at the first clustering the centroids are randomly created. We are now working on the use of different clustering techniques, including hierarchical clustering techniques, implementing the *Clustering* component over Apache Mahout, which implements a Framework MapReduce[4] allowing the usage of different clustering techniques.

In order to validate the quality of the classification of resources obtained, we have evaluated the 12,241 resources considered with the aid of 102 computer science students (advanced and regular internet users) at the Universidad Pública de Navarra during the course 2010-2011. Each reviewer has evaluated a subset of the resource set, ensuring that each resource has been evaluated by five different reviewers. Each reviewer has evaluated its subset of resources after the initial classification, and then those new resources and those whose category of classification has changed along the different recalculations. Reviewers evaluated, for each of the resources, how well resources are classified under their category of classification, quantifying this value between 1 and 5, meaning (1) a very poor classification, (2) a poor classification, (3) a reviewer indecision, (4) a good classification and (5) a very good classification. Reviewers considered 106 resources *very poor* classified (1%), 259 *poor* classified (2%), 4,492 *good* classified and 6,313 *very good* classified. The reviewers were hesitant with 971 resources (8%).

4 Conclusions

We have proposed, implemented and analysed a simple and incremental method for the automatic and semantic creation of concepts to group the resources of a folksonomy, in order to improve the knowledge management in folksonomies, without changing the way users make their annotations. The method automatically creates these concepts and adapts them to the folksonomy evolution over time, grouping new resources and creating, merging or splitting concepts as needed. It is an incremental-aggregation technique that adapts to the folksonomy evolution, without requiring to re-evaluate the whole folksonomy.

The method uses a small subset of tags, the set of more representative tags (S_{rt}), in order to apply it to real folksonomies and their evolution without adversely affecting its performance. Furthermore, the method is based on a component based open architecture. This allows its application to folksonomies with different features and needs, and a useful way to assign names to the classification concepts.

The semantic information assigned to the resource by their annotations allows the automatic classification of the resources of a folksonomy under classification concepts without requiring the intervention of human experts. We have experimentally validated, with the aid of human experts, that the method is able to create automatically these concepts and adapt them to the folksonomy evolution over time, classifying new resources and creating new concepts to represent more accurately the semantic of the resources.

Acknowledgements. Research partially supported by the Spanish Research grants TIN2011-28347-C02-02, TIN2010-17170 and INNPACTO-370000-2010-36. The authors also wish to acknowledge their collaboration to students who have validated the proposal.

References

1. Abbasi, R., Staab, S., Cimiano, P.: Organizing resources on tagging systems using t-org. In: Bridging the Gap between Semantic Web and Web 2.0 (2007)
2. Cattuto, C., Benz, D., Hotho, A., Stumme, G.: Semantic Grounding of Tag Relatedness in Social Bookmarking Systems. In: Sheth, A.P., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K. (eds.) ISWC 2008. LNCS, vol. 5318, pp. 615–631. Springer, Heidelberg (2008)
3. Chi, E.H., Mytkowicz, T.: Understanding navigability of social tagging systems. In: Conference on Human Factors in Computing Systems (2007)
4. Dean, J., Ghemawat, S.: Mapreduce: Simplified data processing on large clusters. Communications of the ACM 51(1), 107–113 (2008)

5. Echarte, F., Astrain, J.J., Córdoba, A., Villadangos, J., Labat, A.: Acoar: a method for the automatic classification of annotated resources. In: Proc. of the 5th Int. Conference on Knowledge Capture, pp. 181–182. ACM, New York (2009)
6. Echarte, F., Astrain, J.J., Córdoba, A., Villadangos, J., Labat, A.: A self-adapting method for knowledge management in collaborative and social tagging systems. In: 6th Int. Conf. on Knowledge Capture, pp. 175–176. ACM, New York (2011)
7. Golder, S.A., Huberman, B.A.: Usage patterns of collaborative tagging systems. Journal of Information Science 32(2), 198–208 (2006)
8. Markines, B., Cattuto, C., Menczer, F., Benz, D., Hotho, A., Stumme, G.: Evaluating similarity measures for emergent semantics of social tagging. In: Proc. of the 18th International Conference on World Wide Web, pp. 641–650. ACM, New York (2009)
9. Mika, P.: Ontologies are us: A unified model of social networks and semantics. Web Semantics: Science, Services and Agents on the World Wide Web 5(1), 5–15 (2007)
10. Peters, I.: Folksonomies: indexing and retrieval in Web 2.0. Knowledge & Information: Studies in Information Science. De Gruyter/Saur (2009)
11. Robu, V., Halpin, H., Shepherd, H.: Emergence of consensus and shared vocabularies in collaborative tagging systems. ACM Transactions on the Web 3(4), 1–34 (2009)
12. Staab, S.: Emergent semantics. IEEE Intelligent Systems 17(1), 78–86 (2002)
13. Steels, L.: The origins of ontologies and communication conventions in multi-agent systems. Autonomous Agents and Multi-Agent Systems 1(2), 169–194 (1998)
14. Yin, Z., Li, R., Mei, Q., Han, J.: Exploring social tagging graph for web object classification. In: Proc. of the 15th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, pp. 957–966. ACM, New York (2009)
15. Zhang, L., Wu, X., Yu, Y.: Emergent Semantics from Folksonomies: A Quantitative Study. In: Spaccapietra, S., Aberer, K., Cudré-Mauroux, P. (eds.) Journal on Data Semantics VI. LNCS, vol. 4090, pp. 168–186. Springer, Heidelberg (2006)

Emerging Concepts between Software Engineering and Knowledge Management

Sandro Javier Bolaños Castro[1], Víctor Hugo Medina García[1], and Rubén González Crespo[2]

[1] District University "Francisco José de Caldas", Bogotá, Colombia
[2] Pontificial University of Salamanca, Madrid, Spain
{sbolanos,vmedina}@udistrital.edu.co,
ruben.gonzalez@upsam.net

Abstract. Software Engineering (SE) uses different theories to empower its practices. One such theory is Knowledge Management (KM), which provides an important conceptual heritage. Our proposal establishes emerging concepts that enrich SE from KM. All these concepts are in between knowledge and software, hence we call them Softknowledge (SK), y Hardknowledge (HK); they constitute Knowledgeware (KW). In this paper we emphasize the intentionality that pertains to these concepts, which is a fundamental characteristic for the development, maintenance, and evolution of software. Additionally, we propose a nurturing environment based on the present proposal.

1 Introduction

Since the term SE was coined [1], there has been constant crisis within this discipline. This crisis can be seen when observing the discouraging figures reported by official bodies like the Standish Group [2], where high percentages of failure on the projects conducted are reported. Aiming at improving such a bleak situation, a variety of software development processes have been proposed, ranging from code and fix [3] to Waterfall [4], Spiral [5], V [6], b [7], RUP [8], among others. On the other hand, methodologies such as XP [9], Scrum [10], Crystal [11], ASD [12] have been proposed to address the development of practices, values and principles [13]. It seems that knowing the way (how), the people (who), the place (where) and the time (when), that contribute to the development, the processes and the methodology around a problem is a suitable roadmap; however, discouraging reports continue to appear.

The present paper points in a different direction, namely knowledge management and it attempts to reduce failure of the software projects. Although this approach was studied already [14], [15], [16], emerging concepts in between the theories of Software Engineering and Knowledge Management have not been proposed to strengthen the understanding and the solutions to some of the most relevant problems encountered in Software Engineering.

L. Uden et al. (Eds.): 7th International Conference on KMO, AISC 172, pp. 13–24.
springerlink.com

2 Problem Statement

Although Software Engineering has been empowered from Knowledge Management, the work done so far has adopted concepts that, from the perspective of Knowledge Management, favor Software Engineering. One of these concepts is what is known as *tacit knowledge* [17]; however, when adopting a concept, not only are the solutions brought in but also the problems associated to the framework within the providing discipline of origin. Even though the classification of both tacit knowledge and explicit knowledge [18] together with its processes framework (SECI) [19] represent a good approximation to many of the problems found in collaborative work – which is typical of software projects such as: the strengthening of knowledge generation [20], knowledge gathering [21], knowledge exchange [22] and knowledge co-creation [23] processes – it is clear that engineers always end up facing the same advantages and disadvantages that have already been identified for the same processes in the field of KM.

The tendency of researchers to narrow the gap between two theories as a means of improving knowledge frameworks causes harmful side effects, which are eventually identified in most cases. There is a special and apparent similarity between software engineering and knowledge management, but you need to check the most favorable concepts carefully and find mechanisms beyond the mere adoption and adaptation of these ideas.

We believe that, more than doing a theory transfer, it is necessary to generate emerging concepts.

3 Knowledgeware: The Missing Link between Knowledge and Software

Software is a product obtained from intellect [24], it is an extension of our thoughts [25]. When software products are made, there are as many variables as people participating in the development process – “Conway’s law” [26], [27]. It could also be claimed that software is knowledge; however, such an statement is too straightforward and it would be reckless not to acknowledge knowledge as intelligence [28], an ability [29], as states of the mind [30], beliefs, commitments, intentions [18] among other features; and although software is pervaded with our reasoning, that reasoning is constrained by the conditions of the programming language and programming paradigms [31], and also by the intended purposes [32] of the documents supporting the software product. The solution to this situation is knowledgeware KW; this new species can be found between knowledge and software, and it takes its traits from both concepts.

Given that, from the perspective of knowledge management, there is a clear distinction between tacit knowledge and explicit knowledge and such a distinction points to the possibility of coding; likewise, in the case of KW, we propose to have a distinction between what we call softknowledge SK and hardknowledge HK. From the point of view of coding, tacit knowledge is orthogonal to explicit knowledge. While tacit knowledge resides in people, explicit knowledge can be stored using physical media, typically IT media. One of the most recurrent concerns is perhaps the way in which organizations gather knowledge and manage it, and so different approaches have been proposed [33], [34], [35], [36], [37], [38]. We consider that leaving tacit knowledge to the world of people is the most suitable approach, our concern should actually focus on the construction of a bridge to join tacit knowledge and software; this bridge is what we regard as SK. Likewise, there should also be a bridge between explicit knowledge and software, which is what we regard as HK.

Knowledge resides in peopleware PW [39], SK and HK reside in KW, and ultimately software resides in hardware. Knowledge management analyses PW from the point of view of the organization; among some of these frameworks we find: organizational knowledge management pillar frameworks [33], intangible asset frameworks [34], intellectual asset model [35], knowledge conversion frameworks [18], knowledge transfer model [36], knowledge management process model [37], and knowledge construction frameworks [38], among others. Regarding these models and frameworks, a lot of effort has gone into defining different types of processes that have an impact on the knowledge of the people who make up the organization. Likewise, within software engineering, the problem of conducting a project has been addressed, where projects themselves are based on organizations consisting of people who must assume the responsibility of a given process in order to obtain software. While in knowledge management processes lead to knowledge, processes in software engineering lead to software. We propose that in order to find a path from knowledge, through KW, up to software, it is necessary to conduct both a traceability process and a representation process. Fig. 1.

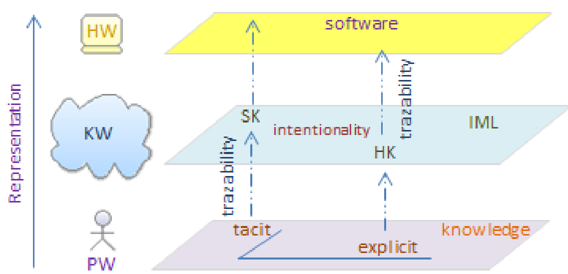

Fig. 1 Knowledge Model

3.1 Traceability

Software traceability [40] is a powerful mechanism that allows the construction of conceptual continuity. Such continuity must permit going back on every step taken in order to understand the origin of concepts [41]. The traceability we propose goes from tacit knowledge to SK, and from there on it goes all the way to end up becoming software; it also goes form explicit knowledge to HK, and form there on it ends up in software.

3.2 Representation

In order to make traceability visible, we propose carrying out representation through languages. If traceability remains over the path tacit-knowledge-to-SK-to-software, we suggest exploiting the advantages of widely-recognized modeling languages, namely archimate [42], UML [43], [44], SPEM [45], or exploiting notations such as BPMN [46] supported by international documents [47]. We also propose the creation of ontological entities supported by languages such as OWL [48]. We believe that software development, when oriented to the creation of models capable of intentional perspectives, can capture certain descriptions where software approximates the knowledge wherein it was created. If traceability also remains on the path explicit-knowledge-HK-software, we suggest exploiting the advantages of having a key between modeling and programming languages, whose perspectives and transformations [49] lead the way from the models language to the programming language.

We coined the concept of Intentional Modeling Language IML, fig. 1, whose additional characteristic over conventional modeling languages will be to provide representation mechanisms that allow expressing the models extended semantics. Some ways to achieve this might be found in the construction of an enriched profiles vocabulary that helps to tinge the models. It is possible to obtain these mechanisms as extension mechanisms in languages like UML. Fortunately, from the perspective of the programming language-transformation-model, this area has been widely developed e.g. MDA [50].

3.3 Intentionality

Most of the problems associated to software lie in the complexity introduced not only by source codes but also by the large volumes of documentation supporting such codes, not to include the typical risks of both coding and documents falling out-of-date. In an attempt to develop and maintain software, engineers often resort to keeping an artifact logbook, where specifications, architectural and design models, user manuals and other records can be found. However, it is by no means an easy task to deal with a product that has been expressed using a language that is intended for a machine as well as using other languages understood by humans. The most considerable difficulty is that the type of knowledge resulting from abstraction, which is expressed in documents and software artifacts, does not easily reflect the actual intention of such creation - of course it is very difficult to capture subjective expressions –; however, it should be possible at least to capture an approximate description. KW was defined as a subjective expression that contains knowledge and that allows itself to be captured through a graphical or textual description by using modeling and programming languages that can incorporate intentional expression mechanisms. Even in art, where so many different emotions are awaken, the piece of art itself provides information about its creation conditions and even about its creator; the piece of art is pervaded by the artist's intentions. Of course in software these intentions must be a lot more bounded; fortunately, disciplines such as software architecture and software design, which leave their mark on software, are very useful to clarify the purpose of how applications are built. In this context, architectural patterns [51], [52], families and styles [53], and also design patterns [54], [55] represent a clearly intentional guide to software.

Other intentional approaches can be used as a mechanism for profiles; in this sense, languages like UML [56] allow an extension of the models semantic meaning through such profiles. Another very valuable approach lies in onthological definitions, which allows adjusting the problem/solution domain with a clear sense of purpose, combating ambiguity. Following these ideas, KW is the traceable and re-presentable knowledge that reflects its intentions in software and therefore facilitates not only software development but software evolution.

4 Intentionality-and-Process-Centered Environment

The importance of processes for both knowledge management and Software Engineering has already been understood, there is even a wide variety of process centred software engineering environments PSEE [57] for software development, which use process modeling languages PML [57]. Nevertheless, although these proposals represent a good development guide when software evolution needs to be evidenced, the roadmaps are not loaded with the necessary intentions to clearly establish the initial purpose of using a given process. We coined the term intentionality-and-process-centered environments.

This type of environment is supported by our own PML proposal. The plus in our PML lies in the SK-and-HK visual modeling from an innovative perspective. Within innovation, we put forward the idea of enriching model semantics and so being capable of building KW. The details of the proposed PML are beyond the scope of this article, but we do illustrate the most relevant concepts from the four points of view of the proposal, namely the knowledge and communications viewpoints, which points to the "software" product; and the improvement and incidents viewpoints, which points to the development process.

4.1 Knowledge Viewpoint

This viewpoint allows modeling knowledge descriptions as well as actors and roles. In this category, a graph permits signaling the most important features from the intellectual assets involved in the development process. Fig. 2 shows the actor called "engineer" with his corresponding role, namely "tester"; the SK is represented as a round-line cloud and HK appears as a straight-line cloud. It can be seen that SK describes ability, while HK describes the possible transformation into a informatics protocol. Both the descriptions are representations of knowledge with a particular intention that allows enriching the semantics of the product.

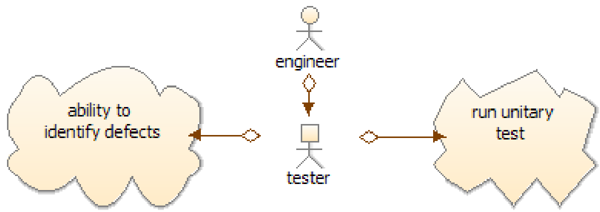

Fig. 2 Knowledge viewpoint

4.2 Communications Viewpoint

Fig. 3 shows how the architect asks the different people involved in the organization about the business nature by conducting an interview whose purpose is to know the architecture. Since the Tower of Babel, the problem of communication has been the obstacle when it comes to setting up projects [24]. Having a simple but expressive vocabulary is another way to achieve clear intentions; we propose modeling a communication interface between actors and roles that expresses the mechanism as well as the ideas arising from a given situation.

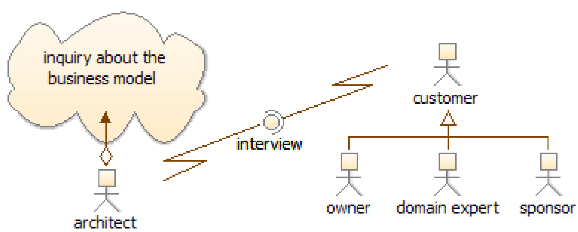

Fig. 3 Communication viewpoint

4.3 Improvement Viewpoint

One of the maxims of software engineering is “process improvement” [58], but yet it has not been directly described in the process model, therefore it is not visible, and because of this, the model and its execution end up being divergent; the person leading the process is supposed to be aware of such an improvement, but a loss of memory is suffered from one process to the other. We propose that the process be enriched with descriptions that point in the direction of improvement. These descriptions can be seen as principles, among other things. Fig. 4.

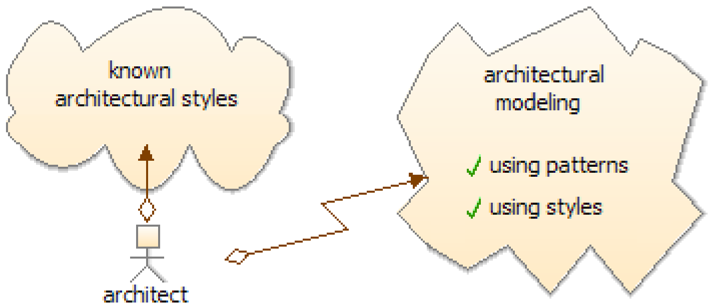

Fig. 4 Improvement viewpoint

4.4 Incidents Viewpoint

In the software development process, another fundamental criterion that must be considered as an intrinsic factor to software is risk [5]; however, like its improvement counterpart, this criterion is not directly associated to a process model either. We believe that risk together with limitations and descriptions are part of a greater set of events that we call incidents, which should be directly reflected in models, Fig. 5.

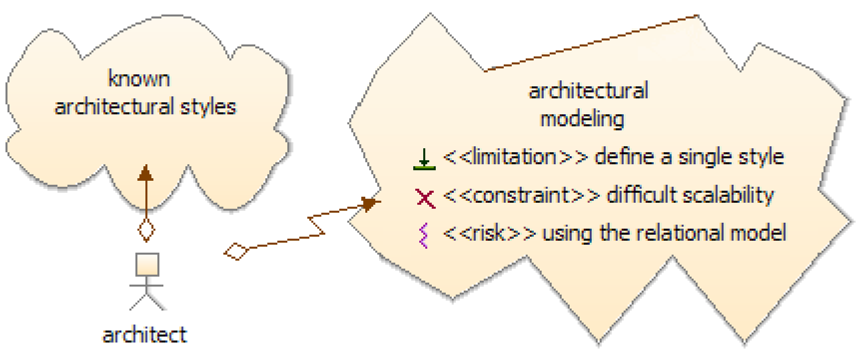

Fig. 5 Incidents viewpoint

The purpose is to provide the models with intentions by taking advantage of either the graphical or textual extensibility mechanisms.

5 Software Tools

There is a wide variety of software development environments that are process-centered, such a Oikos [69], spade [60], Dynamite [61], adele-tempo [62], PADM [63] among others. We propose to have a PSEE plus the models intentions, which we call *Coloso* and will soon be available in its open version at www.colosoft.com.co. This environment integrates Archimate (for the architectural layer) UML (for the design layer) and java (for the programming layer). The whole process is managed from our PML layer. Our proposal also integrates pattern and anti-pattern components, metrics, and process templates. Fig. 6.

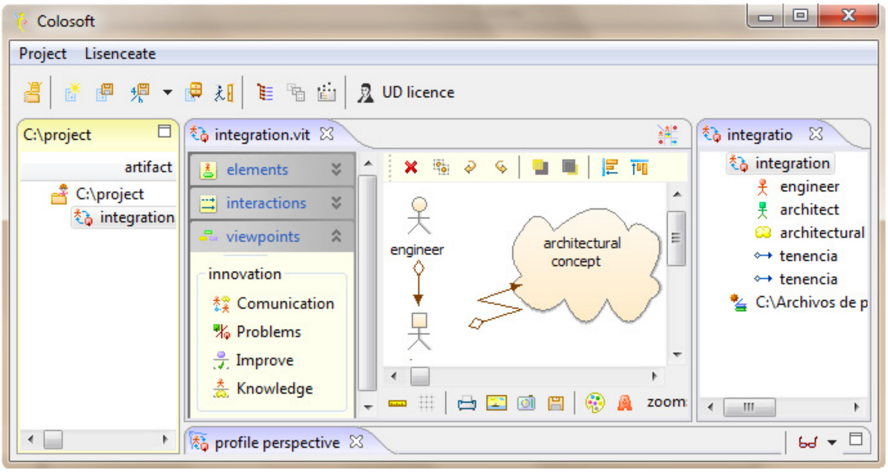

Fig. 6 Coloso

The proposed concept from the *Coloso* platform allows modeling both SK and HK together with their constituent properties. The necessary conceptual continuity between the knowledge layer and the software engineering layer can be established using traceability mechanisms, which are implemented on the platform, including the components that link different types of resources and artifacts such as models, documents, codes and so on. With this approach, it is possible to go from the idea embodied in a conceptual model to its representation expressed in a programming language.

6 Future Work

Three software projects are being developed and tested. An eight-month schedule and a five-people team are the common features of these projects, where two phases are proposed, namely a development phase and a maintenance-and-evolution phase. Our proposal will be used in two of the projects; one of these projects will experience a change in its personnel during the second phase. This event is expected to shed some light on the impact of our proposal.

Another future study addresses the creation and profile-and-ontology enrichment of the *Coloso* environment, which is expected to allow different shadings of the software problem/solution domains. This study is also intended to strengthen the emerging patterns found in between knowledge management and software engineering, which are additional to those patterns already in use [64]; this should allow handling an enriched vocabulary in software processes.

7 Conclusions

Great care is needed when applying theories with the purpose of enriching a discipline; since such theories come together with their own strengths and risks. We believe that the blend of two theories such as SE and KM will necessarily give rise to new concepts like the ones proposed in the present article. The aim of having new concepts is to integrate the strong points; we consider that knowledge is reflected in software, but its intentions will get lost due to the limitations and restrictions that pertain to programming languages unless we decide to use mechanisms like the ones proposed in this article.

Since software is not only about source codes and executable files, we must take advantage of other constructs such as models, which, by using semantic extensions, can be pervaded by the purpose, which makes it easier to perform software development tasks, software maintenance and software evolution. In this article, SK and HK are proposed as two new concepts that can tinge both knowledge and software. By using these two concepts, we propose to construct a bridge between the PW domain knowledge and the hardware domain software. SK and HK must be characterized since they are intentional, that is, they provide the purpose of their very making through different mechanisms, allowing their traceability to be evidenced through the given presentation.

References

1. Naur, P., Randell, B.: Software Engineering: Report of the Working Conference on Software Engineering. NATO Science Committee (1969)
2. Standish Group: The Extreme Chaos. Standish Group International, Inc. (2006)
3. MCconell, S.: Repid Development. Microsoft press (1996)
4. Winston, R.W.: Managing the development of large software systems. In: Proceedings, IEEE Wescon, pp. 1–9 (1970)
5. Boehm, B.: A spiral model of software development and enhancement. SIGSOFT Software Engineering Notes 11(4) (1986)
6. Forsberg, K., Mooz, H.: The relationship of system engineering to the project cycle. In: At NCOSE, Chattanooga, Tennessee (1991)
7. Ruparelia, N.B.: Software development lifecycle models. SIGSOFT Software Engineering Notes 35(3) (2010)
8. Jacobson, I., Booch, G., Rumbaugh, J.: El proceso unificado de desarrollo de software. Addison Wesley (2000)
9. Beck, K.: Embracing change with extreme programming. Computer, 70–77 (1999)
10. Schwaber, K.: Scrum Development Process. In: OOPSLA 1995 Business Object Design and Implementation Workshop. Springer (1995)
11. Cockburn, A.: Crystal Clear. A human-powered methodology for small teams. Agile software development series. Addison-Wesley (2004)
12. Highsmith, J.: Adaptive software development: A collaborative approach to managing complex systems. Dorset House, Nueva York (2000)
13. Manifesto for Agile Software Development, `http://www.agilemanifesto.org/`

14. Ward, J., Aurum, A.: Knowledge management in software engineering describing the process. In: Proceedings of 2004 Australian Software Engineering Conference, pp. 137–146 (2004)
15. Dakhli, S., Ben Chouikha, M.: The knowledge-gap reduction in software engineering. In: Third International Conference on Research Challenges in Information Science, RCIS 2009, pp. 287–294 (2009)
16. Santos, F.S., Moura, H.P.: What is wrong with the software development?: research trends and a new software engineering paradigm. In: OOPSLA 2009: Proceeding of the 24th ACM SIGPLAN Conference Companion on Object Oriented Programming Systems Languages and Applications. ACM (2009)
17. Polanyi, M.: The Tacit Dimension. Routledge, London (1967)
18. Nonaka, I.: A Dynamic Theory of Organisational Knowledge Creation. Organisation Science 5(1) (1994)
19. Nonaka, I., Takeuchi, H.: The Knowledge-Creating Company. Oxford University Press, New York (1995)
20. Les, Z., Les, M.: Shape understanding: knowledge generation and learning. IEEE Transactions on Knowledge and Data Engineering 16(3), 343–353 (2004)
21. Khankasikam, K.: Knowledge capture for Thai word segmentation by using CommonKADS. In: 2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE), vol. 1, pp. 307–311 (2010)
22. Yang, L.: Knowledge, Tacit Knowledge and Tacit Knowledge Sharing: Brief Summary of Theoretical Foundation. In: International Conference on Management and Service Science, MASS 2009, pp. 1–5 (2009)
23. Yao, C.-Y.: Knowledge interaction, adaptive learning, value co-creation and business model innovation. In: Technology Management in the Energy Smart World (PICMET), Proceedings of PICMET 2011, pp. 1–8 (2011)
24. Brooks, F.P.J.: The Mythical Man-Month. Addison-Wesley (1995)
25. Weinberg, G.M.: The psychology of computer programming. Litton educatioal publishing, Inc. (1971)
26. Conway, M.: How Do Committees Invent? Datamation 14(4), 28–31 (1968)
27. Kwan, I., Cataldo, M., Damian, D.: Conway's Law Revisited: The Evidence for a Task-Based Perspective. IEEE Software 29(1), 90–93 (2012)
28. Nelson, R., Winter, S.: An Economic Theory of Evolutionary Change. Belknap Press of Harvard University Press, Cambridge (1982)
29. Wagner, R.K.: Tacit knowledge in everyday intelligence behavior. Journal of Personality and Social Psychology 52(6), 1236–1247 (1987)
30. Schubert, P., Lincke, D., Schmid, B.: A global knowledge medium asa virtual community: The net academy concept. In: The Fourth Americas Conference on Information Systems Proceedings (August 1998)
31. Tucker, A., Noonan, R.: Programming –languages, Principles and Paradigms. Ed. Mc Graw Hill (2002)
32. De Padua Albuquerque Oliveira, A., Do Prado Leite, J.C.S., Cysneiros, L.M., Cappelli, C.: Eliciting Multi-Agent Systems Intentionality: from Language Extended Lexicon to i* Models. In: XXVI International Conference of the Chilean Society of Computer Science, SCCC 2007, pp. 40–49 (2007)
33. Wiig, K.: Knowledge Management Foundations. Schema Press, Arlington (1993)
34. Sveiby, K.: The New Organizational Wealth. Berrett-Koehler, San Francisco (1997)
35. Petrash, G.: Dow's Journey to a Knowledge Value Management Culture. European Management Journal 14(4), 365–373 (1996)

36. Szulanski, G.: Exploring Internal Stickiness: Impediments to the Transfer of Best Practice Within the Firm. Strategic Management Journal 17, 27–43 (1996)
37. Alavi, M.: KPMG Peat Marwick U.S.: One Giant Brain. Harvard Business School (Case), 9-397-108, Rev. (July 11, 1997)
38. Leonard-Barton: Wellsprings of Knowledge. Harvard Business School Press, Boston (1995)
39. Demarco, T., Lister, T.: Peopleware productibe projects dn teams. Dorset House (1999)
40. Seibel, A.: From Software Traceability to Global Model Management and Back Again. In: 2011 15th European Conference on Software Maintenance and Reengineering (CSMR), pp. 381–384 (2011)
41. Bolaños, S., Medina, V., Aguilar, L.J.: Principios para la formalización de Ingeniería de Software. Revista Ingenieria 14(1) (2009)
42. Open Group: Technical Standar Archimate 1.0 Specification. Published by the open group (2009), http://www.archimate.org
43. Booch, G., Rumbaugh, J., Jacobson, I.: The Unified Modelling Language User Guide. Addison-Wesley (2005)
44. Booch, G., Rumbaugh, J., Jacobson, I.: The Unified Modelling Languages Reference Manual. Addison-Wesley (2005)
45. SPEM, Software & Systems Process Engineering Meta-Model Specification. Version 2.0. Final Adopted Specification, http://www.omg.org
46. Object Management Group (OMG): Business ProcessModeling Notation (BPMN), Version 1.2, OMG Document Number: formal/2009-01-03, OMG (January 2009)
47. Mens, K., Mens, T., Wermelinger, M.: Maintaining software through intentional source-code views. In: SEKE 2002: Proceedings of the 14th International Conference on Software Engineering and Knowledge Engineering. ACM (2002)
48. OWL: Web Ontology Language, http://www.w3.org/TR/owl-guide/
49. Zorzan, F., Riesco, D.: Transformation in QVT of Software Development Process based on SPEM to Workflows. Latin America Transactions, IEEE (Revista IEEE America Latina) 6(7), 655–660 (2008)
50. Kleppe, A., Warmer, J., Bast, W.: MDA explained the model driven architecture: practice and promise. Addison-Wesley (2003)
51. Buschmann, F., et al.: Pattern-Oriented Software Architecture a System of Pattern. Wiley (1996)
52. Dikel, D., Kane, D., Wilson, J.: Software Architecture Organizational Pinciples and Patterns. Pretince Hall (2001)
53. Shaw, M., Garlan, D.: Software Architecture. Prentice Hall (1996)
54. Gamma, E., Helm, R., Jonson, R., Vlissides, J.: Design Patterns Elements of Reusable Object-Orienyted Software. Addison-Wesley (1995)
55. Kuchana, P.: Software Architecture Design Patterns in Java. Auerbach (2004)
56. OMG, UML Profile for Schedulability, Performance, and Time Specification (September 2003)
57. Derniame, J.C., Kaba, B.A., Wastell, D.: Software process: principles, methodology, and technology. Springer, Heidelberg (1999)
58. CMMI Product Team, "Improving processes for better services–CMMI® for Services, Version 1.2", technical report, CMU/SEI-2009-TR-001, ESC-TR-2009-001 (February 2009)

59. Montangero, C., Ambriola, V.: OIKOS: Constructing Process-Centred SDEs. In: Finkelstein, A., Kramer, J., Nuseibeh, B. (eds.) Software Process Modelling and Technology, pp. 335–353. Research Studies Press, Taunton (1994)
60. Bandinelli, S., Fuggetta, A., Ghezzi, C., Lavazza, L.: SPADE: An Environment for Software Process Analysis, Design, and Enactment. In: Finkelstein, A., Kramer, J., Nuseibeh, B. (eds.) Software Process Modelling and Technology, pp. 223–247. Research Studies Press, Taunton (1994)
61. Heiman, P., Joeris, G., Krapp, C.A., Westfechtel, B.: DYNAMITE: Dynamic Task Nets for Software Process Management. In: Proceedings of the 18th International Conference on Software Engineering, pp. 331–341. IEEE Computer Press, Berlin (1996)
62. Belkhatir, N., Estublier, J., Melo, W.: ADELE-TEMPO: An Environment to Support Process Modelling and Enaction. In: Finkelstein, A., Kramer, J., Nuseibeh, B. (eds.) Software Process Modelling and Technology, pp. 187–222. Research Studies Press, Taunton (1994)
63. Bruynooghe, R.F., Greenwood, R.M., Robertson, I., Sa, J., Warboys, B.C.: PADM: Towards a Total Process Modelling System. In: Finkelstein, A., Kramer, J., Nuseibeh, B. (eds.) Software Process Modelling and Technology, pp. 293–334. Research Studies Press, Taunton (1994)
64. Bolaños, S., Medina, V.: Knowledge management patterns applied to the software development process. In: 6th International Conference, KMO (2011)

An Ecosystem Approach to Knowledge Management

Vanessa Chang[1] and Amanda Tan[2]

[1] School of Information Systems, Curtin University of Technology,
Perth, Western Australia
vanessa.chang@curtin.edu.au
[2] Deloitte Consulting, Perth, Western Australia
amtan@deloitte.com.au

Abstract. Effective use of knowledge in organisations would typically lead to improved organizational performance. Organisations that practice and leverage organizational knowledge through integration, innovation and sharing of lessons learned will continue to improve and strengthen the organisation's operation. The ecosystem approach described in this paper centred on the premise that knowledge is individualised information enriched through the process of learning, then shared and applied to practical situations. To capture better knowledge management, a collaborative knowledge management framework is used to examine the positions of knowledge management in an organisation. To highlight the complex and heterogeneous nature of knowledge management, a set of practices aimed to enhance collaboration is presented as holistic approach toward improving practical learning environments. In addition, a set of actionable knowledge management strategies to improve the relationships among the components interacting within each organisational ecosystem can be recommended as a result of using the framework.

Keywords: Knowledge Management, Organisational Learning, Learning Ecosystem, Collaborative Learning.

1 Introduction

Today's enterprises are knowledge-based entities where proactive knowledge management (KM) is vital to attaining competitiveness [15]. Nonaka and Takeuchi [22] emphasised that knowledge is a strategic corporate asset that has become "the resource, rather than [just] a resource". Consequently, KM is an integral business function as organizations [3] begin to realise that their "competitiveness hinges on [the] effective management of intellectual resources" [12].

Por [27] articulates that "knowledge exists in ecosystems, in which information, ideas and inspiration cross-fertilise and feed one another", implying that the interactions that occur among individuals, organizations and knowledge artefacts are the primary sources of learning, knowledge creation, sharing and utilisation. Drawing on this insight, this research emphasises that KM is a set of principles

L. Uden et al. (Eds.): 7th International Conference on KMO, AISC 172, pp. 25–35.
springerlink.com

and practices that aims to improve collaborative and cooperative interactions that occur within a particular organizational environment. Janz and Prasarnphanich [16] conducted a study where knowledge is intrinsically gained through cooperative learning with a view that successful organizational KM involves effectively facilitating the ongoing creation, transfer and application of knowledge [2] by creating an environment that is conducive to collaborative learning. This way, learning is direct interactions between individuals and peers in small groups. This environment stresses the importance of just-in-time socialisation or networking in face-to-face or online modes.

Organisational knowledge arises through a continuous dialogue between explicit and tacit knowledge where knowledge is developed through "unique patterns of interaction between technologies, techniques and people" [4]. Individuals continue to develop new knowledge and organizations play a critical role in articulating and amplifying that knowledge [21]. The successful codification of individuals' tacit knowledge to permanently retain as organisational knowledge relies on the effective implementation of KM.

While there has been considerable development in the areas of knowledge, KM, communities of practice and organizational learning, much of the KM research has been accomplished without substantial impact on the way organizations operate [12]. To bridge the gap between theory and practice, this study draws on Brown's [5] proposition that "an organization is a knowledge ecology" comprise of interacting components. The research advances Chang and Guetl's [6] ecosystem-based framework by drawing on the relationship between KM and collaborative organizational learning, and highlights three key categories of contributors to the dynamics of collaborative learning environments; these being (1) the learning or knowledge management utilities, (2) the learning stakeholders and (3) the internal and external influences that could impact the learning environment [6]. The Collaborative Learning Ecosystem (CLES) serves as a valuable framework for examining and improving the KM positions of organizations.

2 Knowledge Management and Collaborative Organizational Learning

McAdam and Reid [19] and Janz and Prasarnphanich [16] observed that knowledge and learning are inextricably related, each relying on and influencing the other. A learning organization is one creates and shares new knowledge, recognizing that its ability to solve problems does not rely solely on technology [12] or in the individual expertise of personnel, but rather, is a result of an interaction among components within the organizational knowledge base [9]. Alavi and Leidner [2] reinforce that bringing individuals together in a collaborative environment to enhanceboth tacit and explicit knowledge [30] is a KM imperative.

Nonaka [21] asserts that "organisational knowledge creation is a never-ending process that upgrades itself continuously"; and new concepts evolve after it is created, justified and modelled. This interactive process that occurs both intra-organizationally and inter-organizationally is facilitated by collaborative organizational learning where the state of knowledge is continually changed to enable its application to a problem or situation [11].

It is clear that effective organizational learning is dependent on the successful creation, storage, retrieval and transfer of knowledge for appropriate application. Learning, knowledge creation and sharing take place at the individual, team and organizational levels, and these processes are vital to organizational learning [11, 2]. The key focus of KM is to encourage and simplify the collaborative learning process into consideration a range of technological, cultural and social factors.

2.1 Development of a Technological Infrastructure

Information technology (IT) serve as an effective enabler [16] of various facets of KM, from the capturing of tacit or explicit knowledge to its application [29]. KM tools assist in the development of K systems involving a combination of people, technology and culture [18].

Without a stable technological infrastructure, an organization will have difficulty enabling its knowledge workers to collaborate on a large scale [13]. The selection of technology, implementation approach and delivery of content must be performed with a focus on users' needs [13]. The ongoing strategic planning and maintenance of IT applications, including ongoing IT risk management that incorporates security and authentication controls [29] and business continuity management, are integral aspects of developing a stable IT infrastructure. Users must be guided on how to use the technology to communicate and share knowledge [13]. Also, if KM is to be a strategic asset rather than a 'passing fad', it must be aligned with economic value which is attainable by "grounding KM within the context of the business strategy" [34].

2.2 A Knowledge-Friendly Culture

Organisational culture is considered the most significant input to effective KM and organizational learning in that culture shapes the values and beliefs that could encourage or impede knowledge creation, sharing and decision-making [16]. Janz and Prasarnphanich [16] suggest that a positive culture is one that promotes knowledge-related activities by giving workers the support and incentives to create an environment that favours knowledge exchange and accessibility. Alavi, Kayworth and Leidner [1] conclude that there exists a positive relationship between a 'good' knowledge culture (defined by trust, collaboration and learning) and a firm's ability to effectively manage knowledge. They express that a culture of trust and collaboration establishes a greater willingness among employees to share insights and expertise. In contrast, "value systems that emphasise individual power and competition among firm members will lead to knowledge hoarding behaviours" [1].

2.3 Communities of Practice and Social Networks

A community of practice is a "self-organised, self-directed group of people" [8], who are held together by a common purpose, shared interests and varied affiliations [5]. These communities of practice gain access to the creativity and resourcefulness of informal groups of peers, allowing its members to "initiate and contribute to projects across organizational boundaries" [27]. These platforms for interaction contribute to the amplification and development of new knowledge [21], as ideas initially formed in individuals' minds are able to be developed through the interactions based on a "passion for a joint enterprise" [8]. Therefore, communities of practice, formal and informal networks are increasingly valued as fundamental platforms for collaborative learning [9]. In today's web-based era, organizations that create an environment that supports the formation of virtual communities of practice or communities of influence can also gain significant benefits in the areas of knowledge transfer, response times and in-novation [8].

Developing a climate conducive to learning is expected to enhance organizational learning, and in turn, business performance [16]. Nunes et al. [23] state that "KM advantages have to be clear and easily attainable; otherwise organizations will continue to focus on the traditional way of working". Earl [10] and Grover and Davenport [12] assert that even if an organisation embraces the concept that well-managed knowledge could enhance performance, they often do not know how to plan and execute KM initiatives. To address the KM challenge where the application of KM is problematic and the factors requiring consideration are often ambiguous, the Collaborative Learning Ecosystem (CLES) framework has been developed to holistically examine unique organizational learning environments.

3 The Collaborative Learning Ecosystem (CLES) Framework

A vital aspect of this framework is the ability to capture the evolving nature of knowledge and hence the term 'ecosystem' was used. The term 'ecosystem' was originally defined by A.G. Tansley as "a biotic community or assemblage and its associated physical environment in a specific place" [26]. The definition implicitly highlights the existence of interactions among the biotic (living) and a-biotic (non-living) components, as well as intrinsically within various highly-complex elements. Pickett and Cadenasso's [26] insights on the applicability of the ecosystem concept to "any system of biotic and a-biotic components interacting in a particular spatial area" led Chang and Guetl [6] to apply the concept to the learning domain in developing an initial "Learning Ecosystem" (LES) framework.

Papows [25] notes that "effective KM systems enable tacit and explicit knowledge to feed off of one another in an iterative manner". It is through this collaboration among the ecosystem components that organisational knowledge is able to grow and be translated into increased value. With a view to highlight knowledge worker as both 'teacher' and 'learner' engaged in collaborative learning activities, the cur-rent research extends Chang and Guetl's [6] framework to incorporate a focus on collaboration. The model is re-named a 'Collaborative Learning Ecosystem' (CLES) to represent the framework's intended application focus.

3.1 Overview of the CLES Framework

The CLES framework (Fig. 1) emphasises "a holistic approach that highlights the significance of each component, their behaviour, relationship and interactions, as well as the environmental borders in order to examine an existing system or form an effective and successful system" [6].

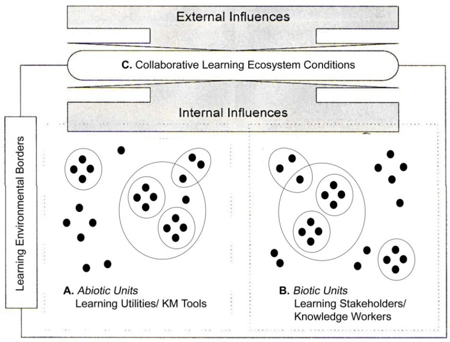

Fig. 1 Representation of the CLES (Adapted from Chang & Guetl [6])

The CLES attempts to simultaneously highlight three key components of contributors (as described below) that "consists of the stakeholders incorporating the whole chain of the [collaborative] learning processes, the learning utilities and the learning environment, within specific boundaries, called environmental borders".

A: Learning Utilities or KM Tools

These are the static and dynamic media that contain and deliver the learning content, and are represented by individual and clusters (each dot denotes a single unit and a group of units make up a cluster that function and work together to serve a purpose) of 'a-biotic units'. These typically (but not necessarily) technology-oriented utilities include all hardware, software and any other computerised platforms that carry the content to the learner.

B: Learning Stakeholders or Knowledge Workers

Learning Stakeholders comprise (i) the learning communities and (ii) other stakeholders who contribute to and/ or benefit from the ecosystem. Learning communities constitute individuals or workgroups (as denoted by clusters of 'biotic units') who can "interact and collaborate synchronously and asynchronously with one another" [6]. Other stakeholders are those who provide the learning content, or support the learning processes through the provision of expertise and services.

C: Collaborative Learning Environment (Restricted CLES Conditions)

The learning environment is dynamic due to changes in a range of internal and external influences, and the impacts of these influences are dependent on the lifecycle of the examined system. External influences include economic dynamics, domain knowledge, competition and technology advancements [24]. Cultural and sociological influences, funding, business strategies and management support are examples of internal influences [6].

3.2 Investigating the Specifics of the Learning Utilities or KM Tools (A)

It is crucial that both IT and non-IT mediums are considered when investigating the range of learning utilities employed by a firm. The ways in which learning utilities are used by stakeholders in carrying out tasks should be examined and where appropriate, represented using Alavi and Leidner's [2] framework shown in Table 1. Once the complete set of learning utilities have been identified, the relationships, usage and implementation effectiveness of each can be investigated. The nature of the interactions among individuals and these utilities is also an issue of interest of the CLES framework.

Table 1 KM Processes and the Potential Role of ICT (Adapted from Alavi & Leidner [1])

<table>
<tr><th>KM Processes</th><th>Knowledge Creation</th><th>Knowledge Storage/Retrieval</th><th>Knowledge Transfer</th><th>Knowledge Application</th></tr>
<tr><td>Supporting Information Technologies</td><td>- Data Mining tools
- Learning Tools</td><td>- E-Bulletin Boards
- Repositories
- Databases
- Search and Retrieval Tools</td><td>- E-Bulletin Boards
- Discussion Forums
- Knowledge Directories</td><td>- Expert Systems
- Workflow Systems</td></tr>
<tr><td rowspan="3">Platform Technologies</td><td colspan="4">Groupware and Communication Technologies
- Instant Messaging, Email, Online Forums, Web 2.0 tools</td></tr>
<tr><td colspan="4">Intranets/ Extranets</td></tr>
<tr><td colspan="4">Internet</td></tr>
</table>

3.3 Investigating the Specifics of the Learning Stakeholders (B)

Learning communities possess learning attributes, which include unique learning styles, strategies and preferences. The learner's demographics, experience, skills, IT competence, objectives, motivations and needs are also important characteristics. Of the range of learning attributes which could influence collaborative learning, a set of characteristics considered to be central are established based on the following characteristics and existing literature.

- Learning style and preference [4]
- Background experiences and perceptions of personal contribution [4, 12, 16, 17, 20]
- Expectations of contributors and usefulness of information [16, 20]
- Motivation to learn [4, 12]
- Motivation to share knowledge [1, 14, 18, 25, 33]

3.4 Investigating the Internal and External Environmental Influences (C)

The application of CLES involves the investigation of the influences affecting a firm's internal and external operating environments. These influences and their impacts usually fluctuate across the business life cycle; and it is vital that the organization adapts to the conditions prevalent at a particular point in time and take action to facilitate the ongoing interaction among the stakeholders and utilities.

A range of existing frameworks can be usedto examine an organization's internal and external environments. Examples of these frameworks include the SWOT analysis, internal value chain or network analysis [28, 31)], Porter's [28] five forces model and the Political, Economic, Social and Technological model.

3.5 Investigating the Internal Environmental Influences

Cultural, business strategies and management support are examples of internal influences of a firm's KM implementation success. The factors and the corresponding literature are used to evaluate the specifics of each internal influence.

- Management leadership and support [7, 13, 14, 18]
- Existence of a knowledge-friendly culture [7, 13, 16, 18, 20]
- Clear knowledge strategy [7, 18, 34]
- Commitment towards an IT infrastructure [2, 7, 13, 18, 29]

3.6 Investigating the External Environmental Influences

Organizations operate within a broad external environment characterised by a climate that is susceptible to radical change [25]. Chang and Guetl [6] consider the industry, government policies, competition, and technology life cycles as important external influences to a firm's collaborative learning environment. The standard Political, Economic, Social and Technological (PEST) analysis could serve as an effective approach to consider the external environment of an organization.

Competition is a key external influence that has an impact on the learning environment of a firm. Porter's [28] Five Competitive Forces model provides a detailed understanding of the competitive environment and relative position of the firm in its industry. The ways in which competitive forces affect a firm's KM and collaborative learning activities can be evaluated based on this analysis.While the PEST and Porter's Five Forces models may not be necessary for every case, each model provides a comprehensive structure for thoroughly evaluating the range of external factors impacting on an organization's learning conditions.

4 Operationalizing the CLES Framework

The CLES framework is believed to provide a holistic approach to facilitate the development of collaborative learning environments. The key to maintaining a positive environment is to “improve the ecosystem as a whole” [6], which in practice, refers to incorporating user-centric collaborative learning, technological innovation, content and learning design in line with the prevalent environmental conditions. These could result in the development of a range of knowledge management practices to help learners respond to uncertain conditions. It is vital that all the components that contribute to or interact within the CLES are appropriately integrated and that a balance in the utilization of each component is achieved [6].

Fig. 2 provides an example of an organization represented as an ecosystem using the CLES framework. This small business general engineering firm operates two distinct business functions of fabrication and machining. The organization which provides specialized services has 7 staff, 2 of whom are office staff, and it is in the early stages of its business.

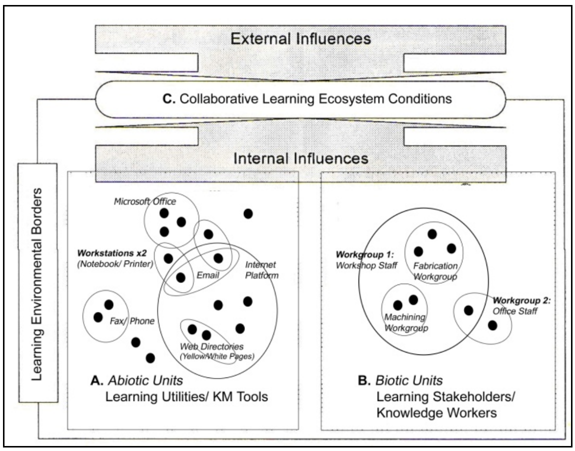

Fig. 2 Representation of a CLES

Analysis was conducted on effectiveness of each learning utility (A), the learning attributes of the learning stakeholders (B) and the levels of interaction among these to carry out work processes. The impacts of various internal and external influences were examined, and based on the analysis; key favourable and unfavourable aspects of the firm’s KM practices were identified.

The favourable aspects were:

1. The users willingly adapt to and make the most of technology-oriented utilities.
2. The information required is usually easily accessible in a readily useable format and does not need to be deciphered.
3. Good use of document templates by office staff to create consistent documents.
4. Frequent use of formal face-to-face interaction platforms to facilitate knowledge transfer and collaborative learning (eg. daily workshop meetings).
5. Active leadership in building the organizational structure and culture.
6. Positive knowledge-friendly culture that is conducive to collaborative learning (staff have a genuine interest in each others’ lives.

7. Minimal barriers to knowledge sharing – on both the social and work levels (atmosphere that values mutual trust, openness and collaboration).
8. Staff have a relatively high motivation to learn and consider their working experience to be a constant learning process.
9. Staff are given opportunities and the flexibility to leverage their backgrounds and experiences.
10. Capable and motivated staff members whose skills are valuable assets.

The unfavourable aspects were:

1. IT utilities are inadequately attuned to users' information needs
   a) lack of integration of IT utilities
   b) risk of human data-entry error due to minimal integrity constraints and in-efficient information sharing and transfer – due to physical separation of workstations.
2. High volume of inaccurate or incomplete documents.
3. Poor (or non-existent) backup and disaster recovery strategies.
4. Inadequate security controls –critical documents are not password protected.
5. Complacent attitudes towards internet security.
6. Significant reliance on individuals' expertise –relies on individuals' 'head knowledge' which is often not documented.
7. Substantial reliance (possibly over-reliance) on ad-hoc (verbal) communication - insufficient enforcement of the use of documents and records for information transfer.
8. Weak records-management and filing procedures.

Based on the above favourable and unfavourable analyses and in view of the apparent lack in technological integration in the firm, the following key recommendations were noted:

1. Actively discover ways to better leverage IT capabilities.
2. Proactively encourage the use of desired learning utilities to gather and accurately communicate work-related information.
3. Implement formal disaster recovery and information security controls.

Based on the premise that the key focus of KM is to encourage organisational learning by taking into account a range of technological, environmental and sociological factors, the CLES framework aims to provide a holistic perspective on the specifics of the (A) Learning Utilities/ Knowledge Management Tools; (B) Learning Stakeholders (individuals and groups of knowledge workers) and the (C) Internal and External Environmental Influences; which contribute to the dynamics of the particular organisation.

The primary aim is to examine the relationships and interactions that occur within the 'environmental borders' of the organisation, and evaluate how effectively KM practices facilitate organisational learning in the firm. Taking into account the behavioural and interaction performance gaps identified, as well as the technological, environmental and sociological factors, a viable set of KM practices to promote more effective collaborative learning is recommended.

In addition to the brief CLES application as provided in Section 5, KM studies using the CLES framework has been carried out on small-to-medium enterprises and these are reported elsewhere [32]. The examination of the firm's KM effectiveness involved a description of its key business processes, followed by an examination of the learning utilities, learning stakeholders and the impacts various environmental influences have on the organization's learning environment. Focus was placed on investigating the range of KM tools and practices employed, and effectiveness of these to support key business processes and organizational value creation; and the impacts of environmental influences on staff interaction, KM initiatives and organizational learning. The CLES allowed a range of significant aspects to be considered in order to make suitable recommendations to facilitate a more conducive collaborative learning environment.

References

1. Alavi, M., Kayworth, T.R., Leidner, D.E.: An Empirical Examination of the Influence of Organisational Culture on Knowledge Management Practices. Journal of Management Information Systems 22(3), 191–224 (2005)
2. Alavi, M., Leidner, D.E.: Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. MIS Quarterly 25(1), 107–136 (2001)
3. Baskerville, R., Dulipovici, A.: The Theoretical Foundations of Knowledge Management. Knowledge Management Research & Practice 4, 83–105 (2006)
4. Bhatt, G.D.: Knowledge Management in Organizations: Examining the Interaction between Technologies, Techniques and People. Journal of Knowledge Management 5(1), 68–74 (2001)
5. Brown, J.S.: Sustaining the Ecology of Knowledge. Leader to Leader Institute (12) (Spring 1999), http://www.johnseelybrown.com/Sustaining_the_Ecology_of_Knowledge.pdf (accessed November 3, 2007)
6. Chang, V., Guetl, C.: E-Learning Ecosystem (ELES) – A Holistic Approach for the Development of more Effective Learning Environment for Small-and-Medium Sized Enterprises (SMEs). In: Inaugural IEEE International Conference on Digital Ecosystems and Technologies (DEST), Cairns, Australia (February 2007)
7. Davenport, T.H., De Long, D.W., Beers, M.C.: Successful Knowledge Management Projects. Sloan Management Review 39(8), 43–58 (1998)
8. Deloitte & Touche LLP (UK) Knowledge Management (2005), http://www.deloitte.com/dtt/article/0,1002,sid%253D5013%2526cid%253D90965,00.html (accessed September 11, 2007)
9. Dickinson, A.: Enhancing Knowledge Management in Enterprises, The Knowledge Board (2002), http://www.knowledgeboard.com/download/593ENKE_KM_Framework.pdf (accessed June 9, 2007)
10. Earl, M.: Knowledge Management Strategies: Toward a Taxonomy. Journal of Management Information Systems 18(1), 215–233 (2001)
11. Gourova, E., Ducatel, K., Gavigan, J., Di Pietrogiacomo, P., Scapolo, F.: Final Report - Enlargement Futures Project - Expert Panel on Technology. Knowledge and Learning (2001), http://forera.jrc.es/documents/eur20118en.pdf (accessed June 7, 2007)
12. Grover, V., Davenport, T.H.: General Perspectives on Knowledge Management: Fostering a Research Agenda. Journal of Management Information Systems 18(1), 5–21 (2001)
13. Hasanali, F.: Critical Success Factors of Knowledge Management (2002), http://www.providersedge.com/docs/km_articles/Critical_success_factors_of_KM.pdf (accessed October 1, 2007)

14. Holsapple, C.W., Joshi, K.D.: An Investigation of Factors that Influence the Management of Knowledge in Organizations. Journal of Strategic Information Systems, 235–261 (2000)
15. Holsapple, C.W., Joshi, K.D.: Knowledge Management: A Threefold Framework. The Information Society 18, 47–64 (2002)
16. Janz, B.D., Prasarnphanich, P.: Understanding the Antecedents of Effective Knowledge Management: The Importance of a Knowledge-Centred Culture. Decision Sciences 34(2), 351–384 (2003)
17. Lee, D.J., Ahn, J.H.: Reward Systems for Intra-Organizational Knowledge Sharing. European Journal of Operational Research 180, 938–956 (2007)
18. Liebowitz, J.: Key Ingredients to the Success of an Orgnization's Knowledge Management Strategy. Knowledge and Process Management 6(1), 37–40 (1999)
19. McAdam, R., Reid, R.: SME and Large Organisation Perceptions of Knowledge Management: Comparisons and Contrasts. Journal of Knowledge Management 5(3), 231–241 (2001)
20. McDermott, R., O'Dell, C.: Overcoming Cultural Barriers to Sharing Knowledge. Journal of Knowledge Management 5(1), 76 (2001)
21. Nonaka, I.: A Dynamic Theory of Organisational Knowledge Creation. Organisational Science 5(1), 14–37 (1994)
22. Nonaka, I., Takeuchi, H.: The Knowledge-Creating Company. Oxford University Press, New York (1995)
23. Nunes, M.B., Annansingh, F., Eaglestone, B., Wakefield, R.: Knowledge Management Issues in Knowledge-Intensive SMEs. Journal of Documentation 62(1), 101–119 (2006)
24. Nyhan, B., Cressey, P., Tomassini, M., Poell, K.: Facing up to the Learning Organisation Challenge: Key Issues from a European Perspective, vol. 1. Office for Official Publications of the European Communities, Luxembourg (2003)
25. Papows, J.: Enterprise.com. Perseus Books Group, USA (1998)
26. Pickett, S.T.A., Cadenasso, M.L.: The Ecosystem as a Multidimensional Concept: Meaning, Model, and Metaphor. Institute of Ecosystem Studies 5, 1–10 (2002)
27. Por, G.: Designing Knowledge Ecosystems for Communities of Practice. In: Advancing Organizational Capability via Knowledge Management. Community Intelligence Labs, LA (1997), `http://www.co-i-l.com/coil/knowledge-garden/dkescop/` (accessed June 7, 2007)
28. Porter, M.E.: Competitive Advantage: Creating and Sustaining Superior Performance. Free Press, New York (1985)
29. Sahasrabudhe, V.: Information Technology in Support of Knowledge Management. In: Srikantaiah, T.K., Koenig, M.E.D. (eds.) Knowledge Management for the Information Professional, pp. 269–276. American Society for Information Science, USA (2001)
30. Sedita, S.R.: Back to "Tribal Fires" Explicit and Tacit Knowledge, Formal and Informal Learning, Towards a New Learning Ecosystem. In: DRUID Summer Conference 2003 on Creating, Sharing and Transferring Knowledge. The role of Geography, Institutions and Organizations, Copenhagen, Denmark (2003)
31. Stabell, C.B., Fjeldstad, O.D.: Configuring Value for Competitive Advantage: On Chains, Shops, and Networks. Strategic Management Journal 19, 413–437 (1998)
32. Tan, A., Chang, V.: An Ecosystem Approach to KM: Case Studies of SMEs. In: 19th Australasian Conference on Information Systems, Christchurch, New Zealand (December 2008)
33. Ward, J., Peppard, J.: Strategic Planning for Information Systems, 3rd edn. John Wiley & Sons Ltd., Chichester (2005)
34. Zack, M.H.: Developing a Knowledge Strategy. California Management Review 41(3) (1999)

Discourse and Knowledge Matters: Can Knowledge Management Be Saved?

Lesley Crane, David Longbottom, and Richard Self

Faculty of Business, Law and Computing, University of Derby
L.Crane1@unimail.derby.ac.uk

Abstract. The Knowledge Management (KM) literature is reviewed with a focus on theory, finding a core issue in the lack of a widely accepted and understood definition of knowledge. Theories are categorised on the bisecting continua of personal vs. organizational knowledge, and reified knowledge vs. knowledge as social action. It is argued that a fresh approach based on the Discourse Psychology framework, and its research tool of discourse analysis, would shed new light on the primary issues. Social interaction – and therefore, language – is considered by many KM theorists to be essential to knowledge sharing and creation, yet language has not been the locus of investigation. DP views language as the site of social action, and reality construction. Consequently, a study of talk in interaction is likely to reveal more about the nature of knowledge and in particular its psychological formulation, with implications for its management.

Keywords: Knowledge Management, Discourse Psychology, discourse analysis.

1 Introduction: A Sea of Theory

There can be few domains of practice and academic inquiry which attract such a broad spectrum of theories and points of debate as Knowledge Management (KM). (Despres and Chauvel, 2002) suggest there are too many theories and that this, combined with what (McFarlane, 2011) calls a lack of a unifying framework on the near horizon, presents a challenge for KM and its practitioners. As to be expected, this patchwork quilt of theories is indicative of definitional difficulties, particularly that of knowledge itself (Quintane et al., 2011). (Bouthillier and Shearer, 2002), like many before and since, construct this definitional problem as a major issue, casting KM as an ill-defined field which none-the-less makes substantial claims. Is KM at risk of drowning in its own sea of theory?

It is also not surprising that a number of studies and commentators point to the high failure rates associated with KM initiatives in organizations (Weber, 2007; Burford et al., 2011). Mainstream KM theory and practice are criticised for relying on flawed or misunderstood assumptions and structures (Virtanen, 2011), while others question what it is that KM is actually managing (Crane, 2011). Most mainstream KM theories set out to design a formula for storing, codifying, transferring

L. Uden et al. (Eds.): 7th International Conference on KMO, AISC 172, pp. 37–46.
springerlink.com

and sharing, and creating knowledge within organizations. The claim is that such a harnessing of knowledge will lead to innovation and competitive edge. However, the risk is that such an approach leads to a reification of knowledge as object (Thompson and Walsham, 2004), and that consequently "real" knowledge is simply devalued and debased. If KM is to evolve and to become an embedded organizational practice, it may be time for a different approach.

This paper argues the case for a fresh theoretical approach to KM. It is proposed that Discourse Psychology (DP) can provide both the theoretical framework, and research methodology, that will pave the way for a different understanding of knowledge and its management in organizations. From the DP perspective, knowledge is not an object to be captured, stored and passed around. Rather, knowledge is something that people do in social interaction: knowledge is constructed and shared in talk and text interaction. Notably, many existing KM theories take this general direction emphasising the concepts of knowledge as action, and the importance of language (Thompson and Walsham, 2004; Blackler, 1993; Burford et al., 2011). DP takes these concepts one step further, focusing the enquiry on how knowledge is created / shared in talk interaction, what discursive resources people call upon to achieve this, and with what consequences.

Beginning with a critical review of a sampling of KM theories, the paradigm of Discourse Psychology and its research method of discourse analysis are investigated. The discussion section considers what Discourse Psychology could bring to KM, with conclusions focusing on the question of "where next?"

2 Theories in Knowledge Management

A contemporary classification organises KM theory into the continua of organizational knowledge vs. personal knowledge, and knowledge as object vs. knowledge as social action.

This sampling of theory – which is by no means exhaustive - shows the majority of theories located along the "knowledge as social action" axis, split between a focus on personal knowledge vs. organizational knowledge. A strong advocate of personal knowledge is (Grant, 2002), who is critical of some theories) for their subscription to organizational knowledge. According to Grant, the unit of analysis must be the person. In his "knowledge-based view of the firm", Grant proposes that knowledge is the most strategically important of a firm's resources, and the purpose of firms is to coordinate teams of specialists, and to facilitate the mutual integration of knowledge. In this model, the goal is knowledge integration. Although Grant's work has been interpreted as a theoretical position, Grant questions this, describing it as a set of ideas that emphasises the role and importance of knowledge. None-the-less, according to Grant, it is the knowledge of individual persons that is the valued asset, and this can only be realised by integration through team-work.

Other personal knowledge theories follow the notion of knowledge as social action more closely. (Boisot's, 2002) I-Frame theory of knowledge creation emphasises the importance of learning as the foundation of knowledge creation, arguing that people do not share knowledge. Rather they share information that becomes knowledge once internalised to the individual. In Boisot's model, knowledge is highly personal, and relies on shared repertoires between individuals. As such it leans more towards a cognitive understanding of knowledge, and raises implications such as people's attention span, perception, short and long term memory, aswell as cognitive demand, although these do not feature in Boisot's framework. The theme that is evident here is that theories in the personal knowledge – social action axis are walking in territory that is of intense interest to the Social Psychologist.

(Blackler's, 1993) "knowledge and the theory of organizations", based on a modified version of Activity Theory, also emphases the central role of social learning. This framework straddles the personal vs. organizational knowledge continuum in it could refer to either. Blackler argues that knowledge is performative, not a possession, preferring the term "knowing" over "knowledge". He is critical of the rational-cognitive mainstream approaches to KM in their reification of knowledge, and their assumption of rationality in both organization and individual – which Blackler disputes. (Blackler, 1995) goes onto argue that there is a shift away from knowledge as situated in bodies and routines (embodied and embedded) towards knowledge as situated in brains (embrained), dialogue (encultured) and symbols (encoded). In other words, Blackler argues for and promotes a move away from knowledge as objective, tangible and routinisable, to knowledge as social action.

Moving more towards the organizational end of the personal-organizational knowledge spectrum, Brown and Duguid place communities of practice at the heart of their proposed architecture for organizational knowledge: "(T)the hard work of organizing knowledge is a critical aspect of what firms and other organizations do" (1999: p 28). They argue that knowledge is mostly collective, and that successful communities of practice are generally informal. However, their architecture is largely dominated by themes of organizational command and control. Note, for instance, Brown et al.'s emphasis on the "organizing" work of firms contrasted with (Grant's, 2002) position on the organization as co-ordinator. (Leonard and Sensiper, 2002) also pursue the notion of people sharing knowledge in group work. Leonard et al.'s theory of creative abrasion frames the different backgrounds, skills, experiences and understood social norms amongst individuals as the factors which generate the melting pot of innovation. In this implied chaotic environment, people will challenge each other leading to an abrasion of different ideas, which in turn gives rise to new ones. As a basic idea, this is not divorced entirely from the concepts of social learning advocated by (Blackler, 1993, 1995) and (Boisot, 2002).

Turning to the other end of the spectrum, the majority of theories taking the line of reifying knowledge largely cluster around a focus on organizational knowledge. The leading theory is that of the knowledge creating firm (Nonaka, 1994; Nonaka and Toyama, 2007). This introduces the SECI model, a dynamic model which explains knowledge creation as an interaction between subjectivities and objectivities. According to this model, new knowledge is created in a spiral of interaction between the processes of socialisation, externalisation (explicit knowledge), combination and internalisation (tacit knowledge).

Similarly to Brown and (Duguid, 1999), (Nonaka, 1994) describes the "informal community" as the location of emerging knowledge, but then suggests that these need to be related to the formal hierarchical structure of the organization. This implies a transformation of the informal to the formal. In his definition of the structure of knowledge, Nonaka proposes that tacit knowledge refers to future events, while explicit knowledge deals with the past, and that only tacit knowledge comprises cognitive elements. There is no evidence for this claim, and arguably, tacit knowledge – even if one takes the view that it is internalised, comprising skills, difficult to articulate, the "more than we can tell" element of knowledge – could refer to the past as well as to the future.

In this brief review of KM theory, I have shown how these are variously focused on personal or organizational knowledge, and vary in their approach to knowledge as either object or as social action. I have also shown how they largely embrace the same factors, but use different terminologies and different emphases. A further commonality across KM theory is the substantial assumptions on which they are based. These include the assumption that knowledge can be identified as a singular thing or activity; that KM outcomes can be measured in some way; that the tacit can be made explicit and vice versa; that knowledge resides in people's heads, but they must be motivated to share it. Others assume that language, communication and social interaction are important, but how is not specified; that what will work in one culture or organization will work in another; and finally, that with the right organizational structure, knowledge can be commanded and controlled.

The paper now turns to consider how Discourse Psychology represents a fresh perspective and approach to KM.

3 Discourse Psychology and Discourse Analysis: New Horizons

Discourse Analysis (DA) is becoming increasingly popular in organizational and inter-organizational research, according to (Phillips and Di Domenico, 2009). As Hardy reports, "(S)such discursive studies are playing a major role in the study of organizations and in shaping some of the key debates that frame organization and management theory," (2001: p 25). There are many different "flavours" of DA (see Phillips et al. for a comprehensive summary, and a review of their application in organization studies), but as Phillips et al. point out, they all share an interest in how social reality is constituted in talk. DA is generally framed within the postmodernist, social constructionist paradigm. This is positioned as diametrically opposed to conventional or traditional theoretical and research approaches. Whilst, as with all theoretical approaches and research methods, DA has attracted its share of criticism, for instance, for the subjective nature of its data (Zajacova, 2002), "(T)there can be no question that the legitimacy of postmodern paradigms is well established and at least equal to the legitimacy of received and conventional paradigms,"(Guba and Lincoln, 2005). The focus here is on the framework offered by Discourse Psychology.

Discourse Psychology (DP) was first introduced by Potter and Wetherell in 1987. This ground-breaking work applies the DP theoretical framework, and its distinctive research methodology of discourse analysis (DA), to the study of attitudes and behaviour as examples of studying psychological phenomena through the lens of discourse. By comparing and contrasting this to conventional theories and methods, Potter et al. highlight the weakness of mainstream approaches, and the advantages of DP in revealing the social world as constituted in language.

The core assumption of DP is that language is the site and location of the social world - human action and performance (as distinct from behaviour). People use language to create versions of the social world. Unlike other brands of DA, the data which DP focuses on includes any form of talk and text, drawn from any medium. Conventional Social Psychology, and research methods in general, approach language as a transparent medium which reflects reality as it is (Marshall, 1994). In the conventional approach, the topic of interest is what people report or say, as a means of uncovering some hidden cognitive structures (Billig, 2001) such as attitudes or beliefs, or intentions to act. Such approaches and methods have come in for considerable criticism over the last two decades (Marshall and Wetherell, 2001; Billig et al., 2003; Antaki, 2000). Rather than study what Billig describes as "…ghostly essences, lying behind and supposedly controlling what can be directly observed" (p 210), if one takes the approach that psychology is constituted in language, then it becomes possible to study the processes of thinking directly.

The assumptions that flow from this basis of language as the site of social action are what make DP singular: that talk is constructive, functional, consequential and variant (Potter and Wetherell, 1987). It is constructive in that talk involves an active selection, which may be conscious or not. This in turn implies that all language is functional in that it works to achieve some accomplishment (e.g., persuasion or argument). It is consequential in the sense that talk construction and function lead to consequences for the speaker and the co-participant(s). It is variable in that one person can describe another, or a phenomenon, action or scene in two completely different ways to two different people.

Unlike traditional approaches, DP and its method of DA are not attempting to understand what is said in talk interaction as a means of shining a torch on some underlying cognitive entity (Potter and Edwards, 2003), rather their focus is on how social reality is produced. (Marshall, 1994) offers a neat and succinct comparison of the relative perspectives on language between the post-modernist and conventional paradigms: traditional approaches view language as a mirror that unproblematically reflects reality, whereas DA approaches language as the site of study in its own right. This juxtaposition of approach immediately highlights a major problem in conventional methods. In its use of questionnaires, scales and so forth, the conventionalist searches for patterns and consistencies in their data with the aim of uncovering phenomena which can be generalised. But, how does the conventionalist accommodate for the variation in language? In this case, the conventionalist must "sort out" any variability in the data as it will impact on results. In the DA paradigm, variability in language is sought for, and studied for its consequence and function. As Potter and (Wetherell, 1987) argue, it is this feature of DP/DA which is its empirical mainstay. Phillips and Di Domenico sum this up: "(L)language and its effects are not a problem to be managed as in positivism, but become the core concern of social science," (2009; p 547).

Thus far, the paper has described the theoretical stance of DP and its research method. Next is a brief review of DA in action, with an overview of the types of insight and knowledge that the discipline has delivered (see Wood and Kroger (2000) for a more extended review). Note that while the following are all examples of DA studies, they are not all strictly in the DP convention.

Identity is a particular area of interest for the discourse analyst. In a classic work, (Abell and Stokoe, 2001)'s analysis of a television interview with Princess Diana shows how she actively constructs two separate identities, one a royal role, the other her private self, exploring the tensions between these, and how she uses her identity work to accomplish consequential actions. In another classic work, Locke and Edwards (2003) study former President Clinton's testimony before the Grand Jury, showing how he constructs his identity as caring, sincere, rational and consistent, and the effects this has on his testimony. A study of courtoom identity demonstrates how identity management can be used to potentially influence a jury (Hobbs, 2003), while another shows how discourse is used to construct the category of "older worker" and its associated meanings (Ainsworth and Hardy, 2004).

Studies of discourse in the workplace include (Alvesson's, 1998) work on how men and women construct gender in an advertising agency context. The findings indicate how identity work by men places stronger emphasis on workplace sexuality in response to the ambiguous nature of advertising work. In a fascinating study of healthworkers' actual practice vs. official policy, Marshall (1994) demonstrates the discrepancies between abstract and actual practice.

Studies of text – including online discussion forums – have also yielded fruitful and meaningful results. Crane (forthcoming), for example, investigates how knowledge practitioners construct identity as expert, creating status, trust and recognition. (Crane, 2011) also investigates how the actors in a forum construct KM, comparing this with issues and debates in the literature. In a persuasive work, (Slocum-Bradley, 2010) analyses discourse in the Danish Euro Referendum, showing how identity construction in discourse could lead to a "yes" or "no" vote.

This is just a tiny sample of the work that has been done. While many of these cases are not situated in the organizational workplace context, their implications for workplace discourse and organizational life are none-the-less valid and transferable. They show how people do the business of constructing everyday realities, and how these realities can often be shown to be at odds with assumed reality. They demonstrate how fundamentally important language is to the action of, for instance, constructing identity, accomplishing persuasion and engendering trust. These are important and influential concepts in the everyday business of the organization. The following discussion section considers some the primary implications that a DP approach might have for KM's primary issues.

4 Discussion: Discourse and Knowledge Matters

To begin, it is worth drawing attention to the fact the majority of studies in the KM domain involve conventional research methods: quantitative surveys, structured interviews and case studies. The introduction of the DP framework represents quite a departure. What will the DP-DA approach bring to KM?

First, consider that the core aim of DA is to shine a light on how people do things in their talk in interaction. Conventional methods of enquiry focus on what people do, with an inferred and subsequent why. By gaining a greater understanding of how people accomplish actions in talk, and with what consequences, the influencing and impacting circumstances can be improved (if needed), fixed (if found to be broken) and applied in ways and to activities not previously countenanced. If it is understood how things are done, these actions and the actions around them can be applied in a more powerful and effective way. To simplify this concept, take (Marshall's, 1994) study of the discourse of health workers. The findings show a discrepancy between official policy and what happens in real practice: while workers advocate and support the "good practice" required of policy, events in reality often prevent or limit such practice. As Marshall argues, such a finding would not have been possible using conventional methods: an analysis of workers' discourse reveals how they construct their work and relationships with others, as opposed to self-reports about what they do (in an ideal world). The problem is revealed, and actions can be taken to address it.

Secondly, the notion of knowledge being shared and created in social interaction with others, and the primacy of language, is a relatively widely shared notion amongst KM theorists and workers. But, language and language practices have not been the focus of enquiry. What DP brings to this agenda is a theoretical framework and a tool for focusing the enquiry on discourse. It is not suggested that, in this sense, DP is a replacement for KM theory, but rather as an extension in a direction which many have already signposted.

Finally, in framing KM's core questions of how is new knowledge created and how is knowledge shared (or how can it be shared) within the DP paradigm, one is immediately forced to consider the nature of knowledge. According to DP, knowledge can be viewed as discursive action, done in social interaction with others: it is, in this view, a social construction. That is not to suggest that an individual cannot create new knowledge on their own: but in this case, individual new knowledge can be seen as the product of their existence in a world of sensory stimulus. This notion of socially constructed knowledge is consistent with some of the KM literature (Blackler, 1993, 1995; Quintane et al., 2011; Burford et al., 2011). It is impossible to avoid asking, if knowledge is action as a social construction, how will we know it when we see it? The position argued here is that knowledge is not a single, well-bounded phenomenon, but rather a collection of psychological phenomena, including, for example, identity construction. If knowledge is approached as a psychological phenomenon or phenomena, then it is appropriate to apply the DP framework to its investigation. It is proposed that the identification of these related phenomena – and they may turn out to be completely integral to what we eventually define as "knowledge" – will lead to a much clearer and accessible understanding of knowledge, and consequently how we do knowledge.

Before turning to some conclusions, it is worth noting a major limitation of the DP-DA approach. The nature of this type of enquiry means that subjectivity – both on the part of the participant and of the researcher – must be carefully attended to (Potter and Wetherell, 1987). The role of the researcher must be included in the focus of the study. While it is important to include context as a core part of the enquiry's focus, it is equally important to manage the experiences, opinions, goals and so forth of the researcher. One method of addressing these issues is to include extracts from the data in study reports to enable the reader to make their own independent judgement of the researcher's interpretation.

5 Conclusions: Where Next?

This paper has attempted to show how KM is an important endeavour for organizations wishing to attain and maintain competitive edge and reap the benefits of innovation. But, the literature indicates there are ongoing, apparently irresolvable issues, predominantly concerned with the definition of knowledge. The paper has shown how KM theories can be categorised into how they treat knowledge: knowledge as object vs knowledge as social action, and personal knowledge vs. organizational knowledge. There is a strong sense that there are too many theories, too many opposing definitions of knowledge, and debates that surround almost all aspects of both theory and practice like a besieging army. It is suggested that a fresh approach is needed, and that this could well take the form of DP-DA. Importantly, DP is not positioned as a replacement for existing KM theories – or their research methods – but rather it constitutes a method and a means for taking these forward along a path that many are already standing at the edge of. Namely, for those who view knowledge as a social construction, with the logical inference of the importance of language, then is it not time to turn the lens of enquiry onto discourse itself as the field in which all of the actions of KM take place? The implications for organizational practice are potentially enormous, and as yet untapped. Anybody need a lifejacket?

References

Abell, J., Stokoe, E.: Broadcasting the royal role: constructing culturally situated identities in the Princess Diana Panorama interview. British Journal of Social Psychology 40, 417–435 (2001)

Ainsworth, S., Hardy, C.: Critical discourse analysis and identity: why bother? Critical Discourse Studies 1(2), 225–259 (2004)

Alvesson, M.: Gender relations and identity at work: a case of masculinitites and femininities in an advertising agency. Human Relations 51(8), 969 (1998)

Antaki, C.: Simulation versus the thing itself: commentary on Markman and Tetlock. British Journal of Social Psychology 39, 327–331 (2000)

Billig, M.: Discursive, Rhetorical and Idealogical messages. In: Wetherell, M., Taylor, S., Yates, S. (eds.) Discourse Theory and Practice: a reader. Sage, London (2001)

Blackler, F.: Knowledge and the Theory of Organizations: Organizations as activity systems and the reframing of management. Journal of Management Studies 30(6), 863–884 (1993)

Blackler, F.: Knowledge, knowledge work and organizations: an overview and interpretation. Organization Studies 16(6), 1021–1046 (1995)

Boisot, M.: The Creation and sharing of knowledge. In: Choo, C., And Bontis, N. (eds.) The Strategic Management of Intellectual Capital and Organizational Management. Oxford University Press, Oxford (2002)

Bouthillier, F., Shearer, K.: Understanding knowledge management and information management: the need for an empirical perspective. Information Research 8(1) (2002)

Brown, J., Duguid, P.: Organizing knowledge. Reflections 1(2), 28–44 (1999)

Burford, S., Kennedy, M., Ferguson, S., Blackman, D.: Discordant theories of strategic management and emergent practice in knowledge-intensive organizations. Journal of Knowledge Management Practice 12(3) (2011)

Crane, L.: What Do Knowledge Managers Manage? Practitioners' discourse in an online forum compared and contrasted with the literature. Journal of Knowledge Management Practice 12(4) (2011)

Crane, L.: Trust me, I'm an expert: identity construction and knowledge sharing. Journal of Knowledge Management (forthcoming)

Despres, C., Chauvel, D.: Knowledge, Context, and the Management of Variation. In: Choo, C., Bontis, N. (eds.) The Strategic Management of Intellectual Capital and Organizational Knowledge. Oxford University Press, Oxford (2002)

Grant, R.: The Knowledge-Based View of the Firm. In: Choo, C., Bontis, N. (eds.) The Strategic Management of Intellectual Capital and Organizational Knowledge. Oxford University Press, Oxford (2002)

Guba, E., Lincoln, Y.: Paradigmatic controversies, contraditions, and emerging confluences. In: Denzin, N., Lincoln, Y. (eds.) The Sage Handbook of Qualitative Research, 3rd edn. Sage, London (2005)

Hardy, C.: Researching organizational discourse. International Studies of Management and Organization 31(3), 25–47 (2001)

Hobbs, P.: "Is that what we're here about?": a lawyer's use of impression management in a closing argument at trial. Discourse & Society 14(3), 273–290 (2003)

Leonard, D., Sensiper, S.: The role of tacit knowledge in group innovation. In: Choo, C., Bontis, N. (eds.) The Strategic Management of Intellectual Capital and Organizational Knowledge. Oxford University Press, Oxford (2002)

Locke, A., Edwards, D.: Bill and Monica: Memory, emotion and normativity in Clinton's Grand Jury Testimony. British Journal of Social Psychology 42, 239–256 (2003)

Marshall, H.: Discourse Analysis in an Occupational Context. In: Cassell, C., Symon, G. (eds.) Qualitative Methods in Organizational Research. Sage, London (1994)

McFarlane, D.: Personal Knowledge Management (PKM): Are we really ready? Journal of Knowledge Management Practice 12(3) (September 2011)

Nonaka, I.: A Dynamic Theory of Organizational Knowledge Creation. Organization Science 5(1), 14–37 (1994)

Nonaka, I., Toyama, R.: Why do firms differ: the theory of the knowledge-creating firm. In: Ichijo, K., Nonaka, I. (eds.) Knowledge Creation and Management: New Challenges for Managers. Oxford University Press, Oxford (2007)

Phillips, N., Di Domenico, M.: Discourse Analysis in Organizational Research: Methods and Debates. In: Buchanan, D., Bryman, A. (eds.) The Sage Handbook of Organizational Research Methods. Sage, London (2009)

Potter, J., Edwards, D.: Sociolinguistics, Cognitivism and Discursive Psychology. International Journal of English Studies 3(1), 93–109 (2003)

Potter, J., Wetherall, M.: Discourse and Social Psychology: beyond attitudes and behaviour. Sage, London (1987)

Quintane, E., Casselman, R., Reiche, S., Nylund, P.: Innovation as a knowledge based outcome. Journal of Knowledge Management 15(6), 928–947 (2011)

Slocum-Bradley, N.: Identity construction in Europe: a discursive approach. International Journal of Theory and Research 10(1), 50–68 (2010)

Thompson, M., Walsham, G.: Placing knowledge management in context. Journal of Management Studies 41(5), 725–747 (2004)

Virtanen, I.: Externalization of tacit knowledge implies a simplified theory of cognition. Journal of Knowledge Management Practice 12(3) (September 2011)

Weber, R.: Addressing Failure Factors in Knowledge Management. The Electronic Journal of Knowledge Management 5(3), 333–346 (2007)

Wetherell, M.: Themes in discourse research: the case of Diana. In: Wetherell, M., Taylor, S., Yates, S. (eds.) Discourse Theory and Practice: A Reader. Sage, London (2001)

Wood, L., Kroger, R.: Doing Discourse Analysis: methods for studying action in talk and text. Sage, London (2000)

Zajacova, A.: The background of discourse analysis: a new paradigm in Social Psychology. Journal of Social Distress and the Homeless 11(1), 25–40 (2002)

An Integrated Pruning Criterion for Ensemble Learning Based on Classification Accuracy and Diversity

Bin Fu[1], Zhihai Wang[1,⋆], Rong Pan[2], Guandong Xu[3], and Peter Dolog[2]

[1] School of Computer and Information Technology
Beijing Jiaotong University, Beijing, 100044, China
{09112072,Zhhwang}@bjtu.edu.cn

[2] Department of Computer Science, Aalborg University, Denmark
{Rpan,Dolog}@cs.aau.dk

[3] School of Engineering & Science, Victoria University, Australia
Guandong.Xu@vu.edu.au

Abstract. Ensemble pruning is an important issue in the field of ensemble learning. Diversity is a key criterion to determine how the pruning process has been done and measure what result has been derived. However, there is few formal definitions of diversity yet. Hence, three important factors that should be further considered while designing a pruning criterion is presented, and then an effective definition of diversity is proposed. The experimental results have validated that the given pruning criterion could single out the subset of classifiers that show better performance in the process of hill-climbing search, compared with other definitions of diversity and other criteria.

Keywords: Ensemble Learning, Classification, Ensemble Pruning, Diversity of Classifiers.

1 Introduction

Ensemble learning refers to a process that learns multiple classifiers to predict the label of an unknown instance, and then combines all the predictions to produce a final prediction. Each classifier learned is also called a base classifier [1]. In the last decade, ensemble learning has gained more and more concerns, due to it could significantly enhance the generalization ability of classification models [2]. In general, the learning process mainly consists of two steps: one is building a number of more diverse and accurate classifiers in training process, the other is aggregating all of the classifiers' predictions to produce a final prediction in classification process [3].

⋆ Corresponding author.

L. Uden et al. (Eds.): 7th International Conference on KMO, AISC 172, pp. 47–58.
springerlink.com