Collective intelligence-based resource allocation to optimise knowledge and innovation harnessing in corporate environments

Ioanna Lykourentzou*

Institut National de Recherche en
Informatique et en Automatique (INRIA),
615 rue du Jardin Botanique,
54600 Villers-lès-Nancy, France
and
Centre de Recherche Public Henri Tudor,
29, Avenue John F. Kennedy
L-1855 Luxembourg-Kirchberg, Luxembourg
E-mail: ioanna.lykourentzou@inria.fr
E-mail: ioanna.lykourentzou@tudor.lu
*Corresponding author

Dimitrios J. Vergados

School of Electrical and Computer Engineering,
National Technical University of Athens,
Zographou Campus, 15773 Athens, Greece
and
Informatics and Computer Technology Department,
Technological Educational Institute of Western Macedonia,
P.O. Box, Fourka Area, 52100 Kastoria, Greece
E-mail: djvergad@telecom.ntua.gr

Amedeo Napoli

Institut National de Recherche en
Informatique et en Automatique (INRIA),
615 rue du Jardin Botanique,
54600 Villers-lès-Nancy, France
E-mail: amedeo.napoli@inria.fr

Abstract: A typical problem that large enterprises face is how to effectively harness the intangible knowledge, expertise, skills and lessons learnt of their staff members. Intangible knowledge harnessing and its codification into usable document formats are vital for the enterprise since they directly affect its ability to innovate and solve complex new problems. However, the identification of the individuals whose contribution can solve each knowledge-demanding problem may be difficult. In this paper, we propose a novel mechanism that uses the collective intelligence of the corporate crowd to identify the tacit knowledge competencies of each employee and coordinate their contributions, inside a wiki-like system, so that each individual may contribute in the most efficient way. Experimental results show that, compared to the fully self-coordinated pattern used by current collaborative knowledge harnessing approaches, the proposed mechanism can help the corporate community allocate its intangible skill resources more efficiently, and thus produce more qualitative knowledge in a timelier manner.

Keywords: collective intelligence; community coordination; resource allocation; corporate tacit knowledge.

Reference to this paper should be made as follows: Lykourentzou, I., Vergados, D.J. and Napoli, A. (xxxx) 'Collective intelligence-based resource allocation to optimise knowledge and innovation harnessing in corporate environments', *Int. J.* , Vol. X, No. Y, pp.000–000.

Biographical notes: Ioanna Lykourentzou is a post-doctoral research fellow at the Institut National de Recherche en Informatique et en Automatique (INRIA) and at the Centre de Recherche Public Henri Tudor. She has worked as a research fellow at the National Technical University of Athens, as well as a Visiting Lecturer at the University of Peloponnese. She received her PhD in Educational and Collective Intelligence System Optimization based on Behavioural Modelling and Machine Learning Techniques from the School of Electrical and Computer Engineering of the National Technical University of Athens (NTUA) in 2009, and her Diploma in Electrical and Computer Engineering from the same school in 2005. Her research interests include the topics of collective intelligence, Web 2.0, corporate collaborative systems and machine learning.

Dimitrios J. Vergados is a Visiting Lecturer at the Technological Educational Institution of Western Macedonia, Kastoria, Greece. He received his Diploma in Electrical and Computer Engineering from the National Technical University of Athens (NTUA) in 2003, and his PhD from the School of Electrical and Computer Engineering of the National Technical University of Athens (NTUA) in 2009. He is also a Researcher at the National Technical University of Athens and at the University of Piraeus. His research interests include design of collective intelligent systems, simulation modelling, wireless networks, scheduling algorithms, and multi-hop networks.

Amedeo Napoli is a Researcher at CNRS since 1984. Currently, he is a CNRS Research Director and the Scientific Head of the Orpailleur team at LORIA-INRIA Nancy Grand-Est. He is a Computer Scientist with a Doctoral in Mathematics and Habilitation in Computer Science. His scientific interests are in the fields of artificial intelligence, and especially in knowledge discovery, formal and relational concept analysis, knowledge representation, reasoning, and semantic web. Among the methods for knowledge discovery, he is a specialist of symbolic methods, such as concept lattice design, lattice-based classification (formal and relational concept analysis), itemset and association rule extraction. Among the methods for knowledge representation, he is a specialist of description logics, object-based representation systems and classification-based reasoning. Recently, he has gained interest in social networks and semantic wikis. Finally, he has been involved in many research projects (international and national) and has authored or co-authored more than a hundred publications.

This paper is a revised and expanded version of a paper entitled 'Collective intelligence based mechanism for the coordination and harnessing of the crowd's innovation potential' presented at 3rd International Conference on COINs, University of Applied Sciences Northwestern Switzerland, HyperWerk Institute for Postindustrial Design, Basel, Switzerland, 8–10 September 2011.

1 Introduction

Corporate innovation, i.e., the firm's ability to generate new knowledge and ideas is acknowledged as a major source of competitive advantage and thus its development is a central target of strategic management (Barney, 2001; Curado, 2006; Sorensen and Stuart, 2000). A key asset for the development of corporate innovation is the effective harnessing of the intangible resources of the corporation, i.e., the tacit knowledge and skills of its individual staff members (Greenwood and Empson, 2003; Kang et al., 2007). This tacit knowledge, expressed as the sum of ideas, lessons learnt and expertise that individual employees have accumulated, is difficult to be codified, compared to the explicit corporate knowledge. Nevertheless, its vital importance drives corporations to continuously seek efficient ways of harnessing and codifying it, as well as ensuring its quality, completeness and reliability.

To perform the harnessing and codification of their tacit knowledge, firms traditionally rely on hierarchical approaches, in which small groups of selected experts collaborate towards the development of new knowledge and solutions under the supervision of a human coordinator (Blankenship and Ruona, 2009; Nickerson and Zenger, 2004). This approach can guarantee the quality and timeliness of the final knowledge result, especially when predefined expertise is necessary (Heckscher and Adler, 2007). However, it often limits the innovation potential of the corporation due to the narrow solution range and the bottlenecks in knowledge production, as a result of the limited number of collaborators and the cognitive saturation of the human coordinator.

Viewing the impressive results often achieved by more participative forms of collaboration in the public internet, corporations have recently started to experiment with crowd-involving approaches, such as prediction markets, crowdsourcing applications and corporate wikis (Surowiecki, 2004).

Prediction markets (Arrow et al., 2008) is an approach that uses combined predictions from the whole employee base to forecast the outcome of an event of particular interest to the organisation. This approach has been found to offer impressive accuracy, which often surpasses that of selected experts, but, on its downsides, it cannot be used for large-scale knowledge development since it can predict only among predefined sets of events.

Crowdsourcing (Chesbrough, 2003), is a second approach, according to which the enterprise outsources an innovation demanding problem not to selected experts but to the whole corporate crowd and rewards the winning solution in monetary terms. This approach is proven to benefit from the crowd's intelligence to produce remarkably novel solutions but, as a drawback, it does not allow for an organisational-wide skill combination, since it is mainly based on competition.

The third crowd-involving approach, which is more suitable for the task of large-scale collaborative knowledge harnessing, is the technology of corporate wikis (Wagner, 2006). Indeed, as recent research shows (Hasan and Pfaff, 2006; Hester, 2010; Lykourentzou et al., 2012; Tapscott and Williams, 2008; Wagner and Schroeder, 2010), the wiki technology is increasingly being used by corporations to facilitate a broad range of knowledge and innovation development tasks, including the intra-departmental codification of tacit knowledge, the formulation of communities of practice, the collaborative development of information systems, the sharing of corporate knowledge with third parties, as well as knowledge preparation for crisis situations. However, current wiki systems, and more generally current collaborative knowledge management approaches, have also been criticised to present serious limitations (Keen, 2007):

questionable content quality, not guaranteed production time (Greenwood and Empson, 2003) and increased overhead costs as a results of the purely self-coordination character of the involved user community (Cao et al., 2009), facts that make the management skeptical into fully employing this technology to the knowledge and innovation development processes of the organisation.

A solution is therefore needed to combine the high quality result that the experts performing in small focused groups can produce with the diversity of ideas and solutions that a collaborating crowd can come up with.

In this paper, and also based on suggestions of the latest research literature, we argue that such a solution lies in reinforcing the current horizontal, crowd-involving approaches, with algorithms that can coordinate the corporate crowd towards producing knowledge content of guaranteed quality, in a timely manner.

To achieve the above, we propose a novel, algorithm-based mechanism that, when applied on a corporate wiki, uses the collective intelligence (CI) of the participating users to identify the knowledge competencies of each individual employee and allocate these to the areas of corporate knowledge that they can improve the most. Specifically, the proposed mechanism uses feed forward neural networks to estimate the tacit knowledge of each individual employee and a resource scheduling algorithm to allocate these competencies in such a way so that the overall goals of the corporation for the production of qualitative and timely tacit knowledge will be met. Experimental results, obtained through the method of organisational simulation, show that the proposed CI mechanism achieves the production of more qualitative wiki articles, in a timelier manner, compared to the respective results achieved by the fully self-coordinated pattern typically observed in current wiki systems. In this way, the proposed CI-based mechanism optimises the performance of the collaborating corporate community and helps it to use its intangible knowledge resources more efficiently.

The contributions of the paper include firstly the proposition of a novel CI-based mechanism that enables organisations to use more effectively the intangible knowledge competencies of their own human resources, and secondly the highlighting of the usefulness that resource scheduling techniques can have on optimising the efficiency of current human ICT-enabled collaboration.

The rest of this paper is organised as follows: Section 2 presents the findings of related literature on the subjects of CI and algorithm-supported mechanisms for the coordination of multiple entities. Section 3 presents the proposed CI-based tacit knowledge harnessing mechanism and highlights the way that it combines the advantages of the hierarchical and horizontal approaches. Section 4 presents and discusses the experimental results, with focus on a specific corporate scenario case. Section 5 discusses the limitations and future directions of this work and finally, Section 6 concludes the paper.

2 Related literature

On the field of algorithm-supported mechanisms to coordinate the collaboration of multiple entities, one may find a number of studies, typically categorised under the field of CI (Levy, 1997; Malone et al., 2009), and especially its computer-science sub-field.

Firstly, a significant number of works focus on the coordination of non-human entities, including multi-agent systems (Wooldridge, 2009), complex adaptive systems

(Holland, 2006) and biomimetic, swarm-intelligence techniques (Engelbrecht, 2006), successfully applied on a variety of problems, ranging from computer network optimisation (Hernández and Blum, 2009) to robot coordination (Pugh and Martinoli, 2009). These studies indicate that the algorithm-facilitated coordination, performed through a distribution of tasks amongst the collaborating entities and a subsequent aggregation of their localised actions, can produce significant collective results, while keeping the individual workload low (Wagner, 2006). However, such CI-based algorithms cannot be directly applied to human communities: careful modification is needed to benefit from the intelligence of the involved human actors, to identify their unique contributions and to remain resilient to their non-deterministic behavioural patterns.

A second type of studies explores the emergence of CI in human communities. The majority of these works use social network analysis to observe the coordination patterns of web communities, to detect closely linked groups within them, to monitor their evolution over time and to measure traits such as trust, consensus-building or group influence (Figueiredo et al., 2009; Kittur et al., 2009; Ortega and Barahona, 2007; Tang and Liu, 2010). The depth of analysis of the above studies is significant and it provides researchers with valuable know-how on the nature and mechanisms that govern web crowd collaboration.

However, in contrast to non-human entity coordination, very few studies use these observations to develop mechanisms which can actively affect the coordination of a human community and systematically help it reach its collective goals. In the following, we highlight the increasing worldwide efforts towards this new direction and analyse the most representative existing works.

A recent research direction uses observations from social network analysis to build mathematical models that can reproduce the collaboration patterns of web communities both for corporate (Lomi and Pattison, 2004) and broader internet (Snijders et al., 2010) contexts, with works on the modelling and simulation of crowd behaviour (Helbing et al., 2011) and network dynamics for, among others, collaborating web groups (Onnela and Reed-Tsochase, 2011).

Another cutting-edge research field focuses on building ensembles of human-machine intelligence that collectively can produce better results than any one of these entities on its own. Works on this field explore, among other CI notions, the potential of using machine intelligence to actively improve the results of crowd collaboration, with applications on collective prediction making (Nagar and Malone, 2011), crowd negotiations (Klein et al., 2007), trend identification and corporate innovation enhancement (Introne et al., 2011).

Specifically on the subject of building algorithms to actively coordinate user collaboration, recent studies have started considering the issue of allocating tasks to users and developing types of crowd-involving recommendation systems. In this context, a number of studies seek to increase collaborative user contributions by matching match users with articles that they might be interested in editing in Wikipedia or in the MovieLens database (Cosley et al., 2007). Other studies use automated text extraction to recommend new content to wiki users and in this way lower the cost of retrieving interesting to update information (Kong et al., 2011). These studies indicate that, similarly to non-human entity coordination, using algorithms to perform task allocation among users can indeed boost the results achieved by collaborating web communities. Nevertheless existing works mainly view the problem from a user-level perspective, seeking to increase individual user participation and in this way to implicitly affect the self-coordination behaviour of the community.

An approach that seeks to achieve not only individual user objectives but also system-level goals set by the overall community could utilise user potential even more.

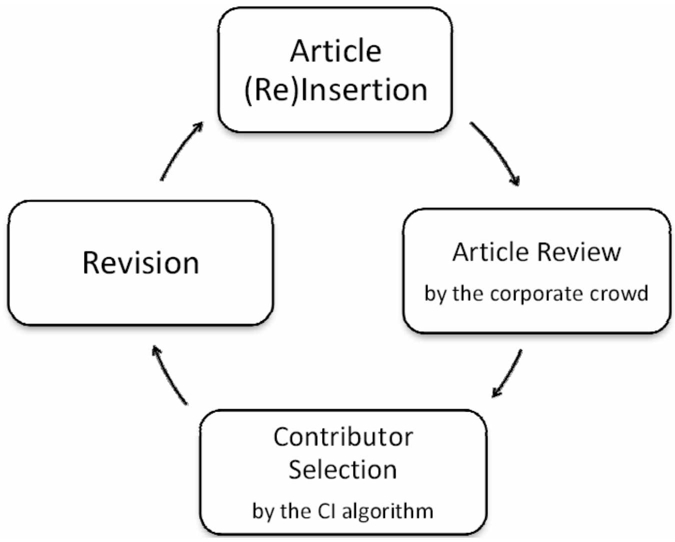
In the context of our previous research (Lykourentzou et al., 2010, 2011b, 2009; Vergados et al., 2010), we have shown that building algorithms to perform this type of *community-wide task allocation* is possible and that it can produce significantly better results, compared to the results of the self-coordination pattern traditionally observed in open user collaboration systems. In this paper, and inline with the aforementioned research findings and trends, we propose a resource allocation, algorithm-based mechanism that uses the CI of the corporate community to perform this type of system-level coordination and in this way boost the innovation potential of the firm.

3 Methodology

3.1 Basic CI-based coordination schema

In the following, we present the basic functionality of the proposed CI-based coordination schema. The target of this schema is to ensure the production of qualitative corporate tacit knowledge and, as a direct result, sustain the innovation production capability of the organisation.

Figure 1 The basic functionality of the CI-based coordination schema consists of successive rounds of community review, selection and revision rounds, until the article reaches its quality threshold



The basic functionality of the proposed schema consists of a series of community-driven insert-review-select-revise rounds (Figure 1). Initially, a piece of knowledge information is inserted into the CI system by a user, in a way similar to the insertion of an article in a wiki platform. Then, a community review process takes place. During this process, the CI system evaluates the quality of the article, based on the opinion of the community. As the literature suggests, fairly accurate quality assessment results can be obtained either through explicit (Lykourentzou et al., 2010) or implicit (Javanmardi et al., 2010) feedback by the community. Based on the obtained quality evaluation, and if the article does not surpass a certain quality threshold, the CI system suggests the article to a selected user for further enhancements. This process of successive rounds of reviews and revisions continues, until the specific information piece is deemed by the community as of satisfactory quality. The algorithm that the CI system uses to suggest articles to users, i.e., the CI-based coordination mechanism, is described in detail in Section 3.2.

- *Combining the advantages of the hierarchical and horizontal approaches:* The proposed CI coordination schema combines the advantages of the horizontal and the hierarchical knowledge and skill harnessing approaches, presented in the introductory section. That is, it uses the corporate crowd's diversity of knowledge and expertise to produce qualitative knowledge articles over a wide variety of topics. In parallel, it retains the benefits of the hierarchical approach, since the CI coordination schema selects the individuals that will be requested to improve each article not randomly, but based on each user's expertise, which the CI system computes as such using the community ratings that the user's past contributions have received.
- *Coordinating the community in a swarm-inspired manner:* Finally, the CI system coordinates the corporate community in a swarm-based manner: each staff member is asked to perform only a few simple actions, like contributing their knowledge on a specific subject, or rating the knowledge contribution of another member without the need of being aware of all the activities taking place simultaneously. It is the algorithms that retain the broader picture, e.g., which knowledge piece needs to be enhanced to meet the organisational demands, as well as a centralised overview of the capabilities of each participant and allocate the available tacit knowledge resources of the community appropriately.

3.2 Resources, tasks, constraints and community skill allocation

In the core of the CI-based coordination mechanism lies the way that the algorithm will decide which member should be requested to revise which article. This can be viewed as a resource allocation problem, with resources, a global goal to be optimised, tasks that the resources should be allocated to and constraints.

Specifically, the *resources* of the global CI system are the capabilities of individual staff members (e.g., individual expertise or individual probability to accept a requested task). The *global goal* that needs to be reached in this case is that each knowledge article should surpass a certain quality threshold. The *tasks* are the review and contributions that each staff members will be requested to perform. Finally, a number of corporate *constraints* can apply, for instance regarding the maximum workload per user allowed, or the number of tasks that a user can decline.

In view of the above, the proposed CI-based coordination mechanism functions in two stages, the resource identification and the resource allocation stage.

3.2.1 Resource identification stage

Firstly, the mechanism estimates individual user expertise, hereby defined as the improvement, in terms of quality, that the contribution of each specific employee would have on each specific wiki article, if that employee contributes their knowledge to that article. As shown in our previous work (Lykourentzou et al., 2010) the estimation of these individual employee resources can be performed using supervised machine learning techniques, such as feedforward neural networks. Machine learning was chosen because, compared to other forecasting methods, it presents the advantage of being data-driven instead of model-driven. Thus, it makes no pre-assumptions regarding the model that governs the future performance of the users, but instead it gradually can 'learn' and adapt to the specific corporate community, the resources of which it needs each time to estimate.

Specifically, using as input the profile of each individual wiki article (consisting of the current quality rating of that specific article and the knowledge area that it belongs) and the profile of each individual user (consisting of the ratings that this user's past contributions have received in the area of the article) we can estimate with satisfactory accuracy (maximum estimation error of approximately 3%) a numerical value representing the quality that the article will have in the future, if the specific employee contributes to it. In this paper we use the above-described resource identification method and focus on exploring the resource allocation problem. For further details on the implementation of this step, and more specifically on the training and validation of the machine learning techniques (minimum necessary training set size, estimation error calculation, etc.) the reader is referenced to the above-mentioned previous work.

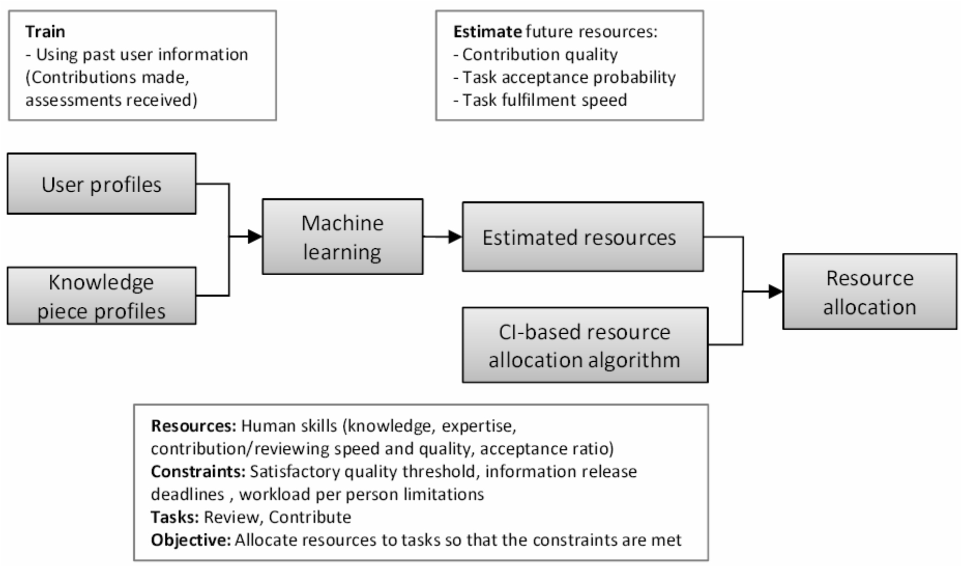
3.2.2 Resource allocation stage

Secondly, using as resources the above estimations regarding the future contribution quality of each employee, and by applying a heuristic resource allocation schema, the mechanism decides which user should be appointed to which wiki article. As with the majority of resource allocation problems, the development of the resource allocation schema to be used greatly depends on the goals and constraints of the specific corporate community, the collaboration of which we each time need to optimise, i.e., on the formulation of the specific resource allocation problem. For example, different allocation schemas can be developed depending on the maximum workload permitted per employee, on the minimum threshold of quality that is considered satisfactory or on the time that this threshold should be reached by each article within the system.

The specific corporate use case scenario and the respective heuristic resource allocation schema, used in this paper to examine the performance of the proposed CI-reinforced approach, are presented in Section 4.

Figure 2 illustrates the above-presented resource identification and allocation stages, which together form the coordination mechanism of the proposed CI reinforced tacit knowledge harnessing approach.

Figure 2 The crowd coordination mechanism of the proposed CI-based tacit knowledge harnessing approach consists of two main stages: the resource identification and the resource allocation one



Given the above, in the following section we present the experimental results that were performed to compare the capabilities of the proposed coordination schema, to the fully self-coordination pattern traditionally observed in current collaborative content creation communities.

4 Results

In this section, we examine the performance of the proposed CI-reinforced tacit knowledge harnessing approach, compared to a *benchmark* system, which represents the functionality of a fully self-coordinated knowledge harnessing system, i.e., a typical corporate wiki system.

The organising of this section is as follows: first we describe the performance metrics that we will use to measure the performance of the proposed approach. Secondly we present the experimental setup set to measure these metrics and last, we present and discuss the obtained results.

4.1 Performance metrics

We will compare the performance of the proposed approach to that of the benchmark system on the basis of the following two metrics:

1 *Quality:* Quality among the most important characteristics of knowledge that the corporate management seeks from the use of any tacit knowledge harnessing approach. Given a target quality threshold set by the management for every produced

article, we will measure the number of articles that surpass this threshold, for both the CI-reinforced and the benchmark systems:

$$Q = \frac{\sum_{k=1}^{n} i_k}{n}, \text{ where } \begin{cases} i_k = 0, & \text{if the quality of } i_k < Th \\ i_k = 1, & \text{if the quality of } i_k > Th \end{cases}$$

where n is the total number of articles that exist in the system at the time that the quality metric Q is measured, i_k is the k^{th} article and Th is the threshold that the articles should surpass to be considered reliable from the corporate management.

2 *Timeliness:* The second metric to evaluate the performance of the proposed approach is timeliness. Given that each article has a deadline pre-defined by the management, i.e., a time period within which it should reach its quality threshold, we will measure the number of articles that are completed on time, for both the CI-reinforced and the benchmark systems:

$$T = \frac{\sum_{k=1}^{n} j_k}{n}, \text{ where } \begin{cases} j_k = 0, & \text{if at time } D \text{ the quality of article } j_k < Th \\ i_k = 1, & \text{if at time } D \text{ the quality of article } j_k > Th \end{cases}$$

where n is the total number of articles that exist in the system at the time that the timeliness metric T is measured, j_k is the k^{th} article, Th is the article's threshold and D is the deadline of the article.

Given the above two metrics, we will measure whether the proposed approach can help the corporate community produce more qualitative articles in a timelier manner.

4.2 Experimental setup

To measure the performance of the proposed and the benchmark systems we used the method of organisational simulation, a process typically adopted by corporations to examine unprecedented systems in a qualitative manner, in order to evaluate their effectiveness prior to a real-world application (Rouse and Boff, 2005; Zacharias et al., 2008).

The simulation modelling and validation was performed in three steps:

1 *Benchmark system modelling:* First we model the benchmark system, which simulates the functionality of a typical corporate wiki system. The benchmark system operates, like typical wikis, according to a fully self-coordinated collaboration pattern where users self-identify and self-select the articles that they will contribute to.

2 *CI-reinforced system modelling:* Secondly we model the CI-reinforced system, which simulates the functionality of the benchmark system with the CI-based coordination schema is applied on it. Users, in the CI-reinforced system, behave in the same way as they would in the benchmark system, but they are also given recommendations by the coordination algorithm regarding the articles that they could contribute to. Therefore, this step also includes modelling of the identification and allocation algorithm.

3 *Performance evaluation:* Finally, we run the simulated systems and compare their performance using the quality and timeliness metrics defined above (Section 4.1).

In the following, we describe in detail the realisation of the above three simulation modelling and validation steps.

4.2.1 Benchmark system modelling

The following elements are important to the modelling of the environment in which the corporate community functions:

Article

An article is a piece of information content inside the collaborative enterprise content creation system, similar to a wiki page. The set of all the articles inside the system is referred to as $A = 1$. Every article inside the system is characterised by a certain quality level q. The quality of an article is modelled as a single numerical value in the [0, 10] range. The initial quality of an article is zero (stub articles). The revision of an article changes its quality. Each article is also characterised by a specific deadline, i.e., the time in which it should, according to the organisational demands, reach a certain quality, modelled following a gamma distribution.

Knowledge domain

The set of articles inside the system are considered to be divisible into a number of domains, with each domain containing a subset of the articles, and representing a specific area of knowledge. The set of all the domains of the system is referred to as $D = \{d_1, d_2, \ldots, d_{|D|}\}$. For simplicity reasons, each article is considered to belong to exactly one domain, that is:

$$\forall d_i, d_j \in D, \quad d_i \cap d_j = \varnothing \tag{1}$$

and

$$\forall a_k, \exists d_j : a_k \in d_j \tag{2}$$

The division of articles to domains could be performed either manually, or through automated content categorisation methods such as the ones presented in Hao et al. (2007).

User

Users represent the participating staff members. Their role is to improve articles through edits. The set of all users inside the system is referred to as $U = \{u_1, u_2, \ldots, u_{|U|}\}$.

Each user u_i is characterised by an expertise vector E_i, with $|E_i| = |D|$ which models their expertise on each one of the knowledge domains. Every element in the expertise vector is given value in the [0, 1] range, with 0 representing no knowledge, whatsoever, of the domain, and 1 representing perfect knowledge. Each value of the expertise vector is initialised independently for each user through a random variable. The expertise of a user in a domain is considered to remain unchanged. In addition, every user is characterised by two different states: 'idle' and 'busy' to indicate that the user is available or engaged with a specific task request.

Finally, for the purposes of this simulation we adopt a non-preemptive approach, i.e., a user cannot be interrupted with a new task once in the 'busy' state.

Benchmark system

Next we model the benchmark system, i.e., the way that the community typically creates its collective tacit knowledge, without the use of the CI coordination schema: users arrive to the system following a Poisson arrival process and they randomly view an article. Taking into account that in collaborate content creation systems, the highly ranked articles tend to attract significantly more attention than lower-ranked ones (Kittur and Kraut, 2008; Zhang et al., 2010), the probability of each user viewing a specific article is modelled proportional to the exponent of the article's quality, i.e.,

$$P\{u_i \text{ views } a_j\} = \frac{b^{q_i}}{\sum_{k=1}^{|A|} b^{q_k}}, \tag{3}$$

where b is set equal to 1.5.

After viewing an article, a user may choose to revise it and therefore change its quality or leave without contributing. Since users are more likely to contribute to a subject that they know about, the user contribution probability to an article is modelled to be proportional to the user's expertise in the domain of the article:

$$P\{u_i \text{ edits } a_j \mid u_i \text{ views } a_j\} = 0.05 e_i^{d_j}, \tag{4}$$

where e_i^{dj} is the corresponding expertise value of user i in the domain d_j of article a_j.

After the edit has been completed, the quality of the article changes taking into account the current quality of the article and the expertise of the user, as follows:

$$q_{new} = 10 \cdot e_i^{d_j} + r \cdot q_{old} - r \cdot e_i^{d_j} \cdot q_{old} \tag{5}$$

The rationale behind this equation is that users with higher expertise improve articles more than non-expert users, and in parallel, articles with low quality improve more rapidly than high-quality ones. It should also be noted that the above equation includes the possibility that a user contribution might worsen an article, according to the reduction factor r, which is set to 0.85. After a successful edit, the above process is repeated.

4.2.2 CI system modelling

The CI system extends the benchmark system, presented above, by suggesting articles to users, with an aim to maximise the community objective, while satisfying the underlying constraints. The modelling and implementation of the CI system and of the coordination algorithm that the system uses depends largely on the specific conditions (e.g., global system target and specific organisational constraints) that apply for the specific organisation according to its demands regarding the production of tacit knowledge. In the following, we present the specific corporate scenario case and the respective CI coordination algorithm modelled in this paper to test the proposed methodology.

The collective community goal, and therefore the *goal of the CI system*, is to release articles of acceptable, to the organisational management, quality. The quality threshold for the present scenario is set to 7, on a scale from 0 to 10. The *system resources* are:

1 the estimated individual user expertise, as estimated by FFNNs as detailed in Section 3.2.1

2 the probability of a user accepting a recommended task

3 the estimated user task fulfilment speed.

All three of these resources are estimated for each user and this estimation can be performed using the technique of feed forward neural networks, as shown in our previous work (Lykourentzou et al., 2010). The *task* that the system should allocate its resources is, for this scenario, the contribution task, i.e., the system selects which user should be requested to improve which article. The *constraints* in this scenario setting are time and workload. Specifically, the system needs to allocate its resources in such a way as to respect the deadline of each article, as well as the fact that each user can be engaged with only one article at a time and that this is a pre-emptive, uninterruptible process. Finally, the *resource allocation algorithm* for this scenario is the following: For every article that has not reached satisfactory quality levels and the deadline of which has not elapsed, the algorithm selects those individual staff members whose contribution can be made on time and ranks them according to their estimated contribution quality. From the pool of these individuals, the algorithm selects the first, i.e., the one that is estimated to be able to improve the article the most and sends this individual the task request. If the specific staff member accepts, the article is allocated to this individual. In case the user rejects, the algorithm moves forward to the next staff member on the list until the end of the list. If no one accepts, the algorithm re-calculates the list and starts over. Table 1 illustrates the above-mentioned CI problem and the heuristic resource allocation algorithm used to address it.

Table 1 The corporate scenario case and the CI-reinforced community coordination algorithm implemented

CI system goal	Individual article quality threshold > 7/10
System resources	• Individual expertise • Individual task acceptance probability • Estimated user task fulfilment speed
Allocation task	Contribution
Constraints	• Time (need to respect each article's deadline) • Max. workload (= 1 article/user, pre-emptive, uninterruptible process)

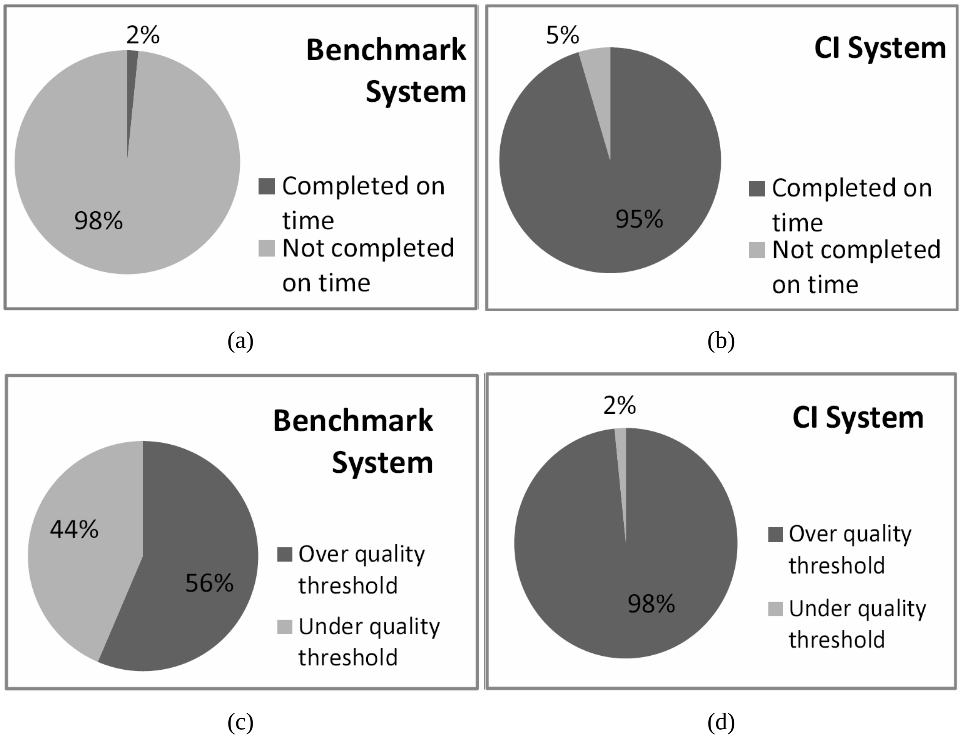
CI-based resource allocation algorithm	For every article not reached threshold + not elapsed deadline:
	{
	1 Select users with estimated timely contribution
	2 Rank them based on estimated contribution quality
	3 Select the first on the list
	4 If user accepts → end;
	5 Else → move to next user on list
	}
	*If no one accepts, the algorithm re-calculates the list and starts over.

4.3 Performance evaluation results

After modelling the benchmark and CI-enabled systems, the simulation validation phase took place. The simulation parameters used are the following: the size of the simulated employee population is 1,000 users, who created, reviewed and edited 1,000 articles. To measure each one of the metrics set, ten simulation rounds took place and the obtained results were then averaged.

As one may observe [Figures 3(a) to 3(d)] the CI coordination schema manages to achieve significantly better allocation of the resources of the community. This is reflected firstly through the number of articles that manage to surpass the desirable quality threshold on time [Figures 3(a) to 3(b)], a number that for the case of the CI schema is significantly higher compared to the respective number of the benchmark system. This result indicates that the community coordination mechanism managed to select those individuals that can complete the articles on time, for the majority of the articles within the system.

Figure 3 The CI-reinforced coordination schema achieves better allocation of the resources of the corporate community in terms of [(a) and (b)] articles reaching the desirable quality thresholds within their deadlines and [(c) and (d)] overall quality levels

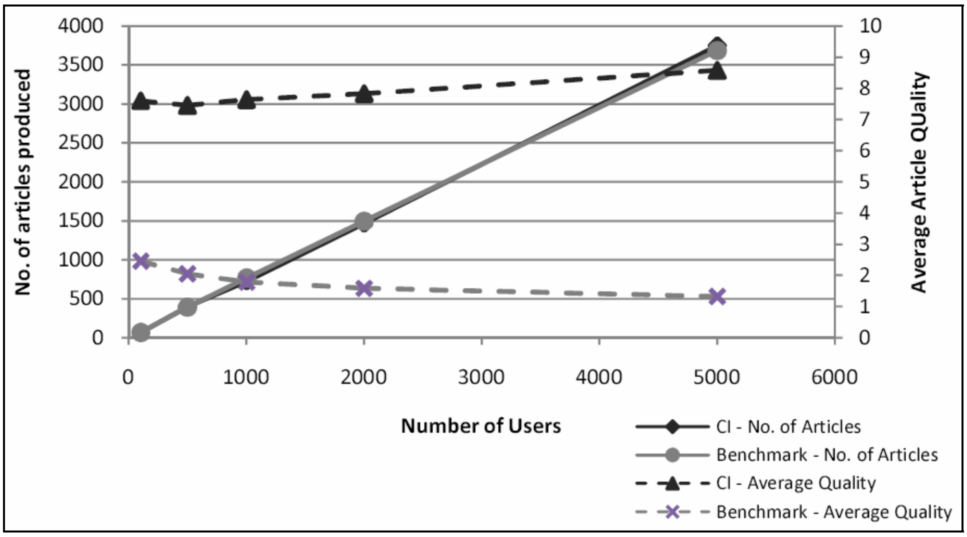


Similarly, the allocation made by the CI coordination schema significantly enhances the overall quality of the articles within the system, since the number of articles that surpass the quality threshold (irrespective of the deadline) in general is significantly higher than the respective ones of the benchmark system.

The above results were obtained for a fixed number of users, corresponding to a medium-sized organisation of 1,000 users. However, it would be also interesting to observe the behaviour of the proposed algorithm on different sizes of knowledge base organisations, ranging from small and medium enterprises (SMEs) (Corso et al., 2003; Tan and Hung, 2006) (up to 500 employees) to large corporations (over 2,000 employees) (Hester, 2010).

Figure 4, illustrates this case, with axis x corresponding to different numbers of participating employees and axis y illustrating the total number of articles produced and the average article quality reached. As it may be observed, the total number of articles increases with the availability of more employees, for both the CI and the benchmark systems, a result which is expected. What is mostly interesting to observe is that at all cases of simulated participating populations the average quality of the articles remains steady, higher for the CI system and lower for the benchmark. This indicates that even with lower numbers of participating employees, the algorithm manages to benefit from the available knowledge competencies, and to allocate these in such a way as to produce articles of satisfactory quality.

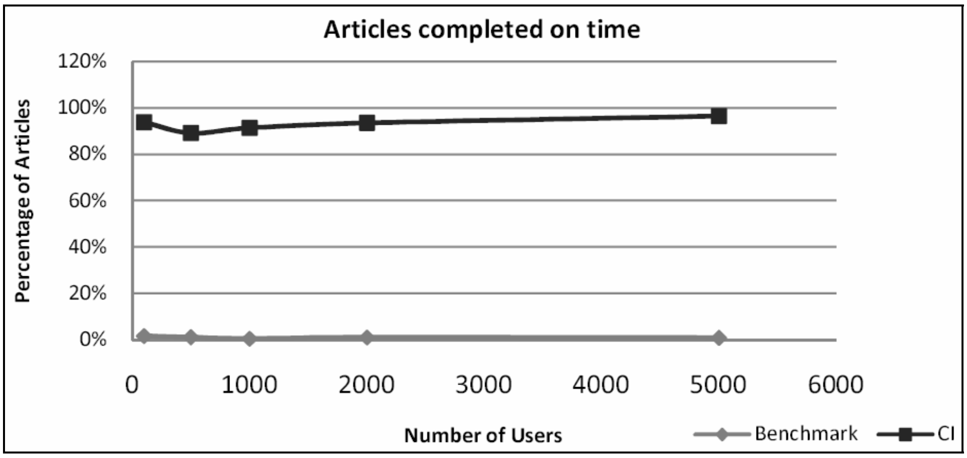
Figure 4 The performance of the CI and benchmark systems (total no. of articles and average article quality) for different knowledge base population sizes (see online version for colours)



Continuing on the experiment of varying population numbers, Figure 5 illustrates the number of articles completed on time (Figure 5) for both the benchmark and the CI algorithms. As it may be observed the algorithm manages to keep a significantly higher number of articles completed on time, compared to the ones achieved through the benchmark system, at all population variations.

Combined, the results of Figures 4 and 5 indicate that the performance of the algorithm remains relatively unchanged with population variations, a fact which in turn indicates that the proposed approach could be beneficial to organisations with different availabilities of employee bases.

Figure 5 Total number of articles completed on time for different knowledge base population sizes



5 Limitations and future work

5.1 Limitations

This paper highlights the potential of enhancing the quality of corporate tacit knowledge, by automatically allocating the knowledge competencies of individual users, inside a wiki, to the articles that they can improve the most.

A number of limitations may be found to the present work. First, the performance evaluation results, as presented in Section 4.3, are obtained through simulation, a technique that on its own implies a certain degree of simplification of the real world. A shift to an actual corporate wiki setting is necessary to evaluate the potential of the proposed approach in improving the quality of organisational knowledge harnessing with more confidence. Such a shift however would come with certain implications, both practical and managerial, which are presented briefly in Section 5.2.

Secondly, even at simulation level, the conclusions drawn on the algorithm performance cannot be generalised. For the time being, we can claim effective algorithm performance only on the specific simulated environment and its parameters. In order to assure effectiveness of the proposed approach it is essential to run significantly more and differentiated scenarios. Certain example scenarios that the authors consider evaluating are also presented in Section 5.2.

Thirdly, the proposed approach presents certain limitations regarding the algorithm implementation itself. The main one refers to the cold-start problem. Currently, the algorithm estimates user contribution quality using neural networks, which require the presence of past user information (e.g., past contributions and how these were rated by other users) in order to be effectively trained and used for future predictions. This in turn introduces a delay in the time that the proposed approach can start being used. Other techniques, such as active learning, can be employed in the place of neural networks, in order to immediately start estimating user competency and then to improve it gradually, as more data become available. In addition to the above, the heuristic resource allocation algorithm used in this paper, to implement the proposed approach, can be ameliorated by shifting the problem directly to the domain of resource allocation (for instance viewing it as a job shop scheduling problem) and applying optimisation techniques from this field to our specific problem of people-to-task allocation. Finally, instead of applying an one-step planning forward algorithm, additional planning forward steps could also be incorporated in the algorithm functionality to enable the corporation make a more efficient use of its employee resources and better prepare for future expected knowledge production loads.

5.2 Future work

Future work mainly includes two directions:

1 developing additional simulation scenarios to further examine the robustness of the proposed approach

2 shifting the proposed approach on real corporate setting.

5.2.1 Future simulation scenarios

On the first direction, three main future simulation scenarios are planned, regarding:

1 *Organisational knowledge sharing culture:* Organisational culture is a factor that directly affects the degree to which employees are willing to share their knowledge. That is, in strictly hierarchically-based organisations, knowledge is considered to be a personal asset and thus it is less likely to be openly shared. On the contrary, more flat-based organisational structures promote the open sharing and exchange of knowledge resources. In the context of our simulated environments, organisational culture would directly affect the probability of a user accepting to contribute to a requested article, as well as the probability of this individual to share their 'true', 'internally-modelled' knowledge. Therefore, as a first future scenario we plan to experiment with variations of these two parameters to examine the performance of the proposed approach in respect to different types of organisational knowledge sharing cultures.

2 *Social network relations among employees:* A second factor to take into account is the social relations, e.g., prior collaborations, that exist between employees of the same organisation. These additional information can then be used to extend the functionality of the proposed approach from fully centralised (where users individually interact with the wiki system) to socially-networked (in which users can suggest reviewers, the system can recommend articles to groups instead of individuals users, etc). This second scenario should therefore model collaborations within the corporate human network, extend the algorithms to include notions such as trust and reputation and compare the obtained performance with that of the approach presented in this paper.

3 *Semantic knowledge domain inter-relations:* The current approach considers a fixed taxonomy, where wiki articles are divided into distinct, non-related knowledge domains. The third future scenario refers to modelling domain inter-relations, which correspond to the semantic interrelations that exist among the knowledge domains of a real-world wiki. This scenario will enable the development of an extended version

of the current algorithm, which will seek to use the domain interrelation information to mitigate the effect of the cold-start problem, in which the knowledge competency of a user in a domain that the user has not yet contributed to can be approximated by the contributions that the user has performed in related domains.

5.2.2 Extension to real-world corporate setting

Apart from examining additional research scenarios, the second direction of future work should include examining the proposed approach on the premises of an actual organisation. Such an evaluation is expected to enable the measurement of elements such as the level of acceptance of the proposed approach and the return of investment, as well as to fine-tune the algorithms developed through simulation to match the needs of the specific corporate community. Shifting the proposed approach to the setting of a real-world wiki would naturally have a number of additional practical and managerial implications (Lykourentzou et al., 2011a; Yates et al., 2010).

Practical challenges include selecting, among a high number of available ones, the appropriate wiki platform, which should be then extended to incorporate the proposed algorithm. The selection of the platform should be driven by the need to balance simplicity and to cover the features necessary for the particular knowledge harnessing task that the wiki will be used. Another practical challenge is the provision of structural support, or in other words the development of a domain taxonomy, in which the wiki articles will be categorised. The development of this taxonomy should be made by the organisational management, in collaboration with domain experts, in regards to the specific knowledge harnessing process that the wiki will facilitate.

Managerial implications are also expected to occur. These mainly include the need to ensure support, both from high and middle-level management, as well as the need to provide participating employees with the appropriate knowledge sharing incentives. Specifically in order for the organisation to benefit from the capabilities of the proposed approach, the management should embrace the need for a non-hierarchical, open and bottom-up knowledge sharing culture, actively urge users to share their knowledge inside the corporation and promote interaction and feedback among the prospective wiki users. Knowledge sharing incentives are also crucial for the success of the proposed system. Specifically, the four main incentives that the management should seek to provide potential users with include:

1 incorporating the wiki to the daily routine of the employees

2 providing recognition for one’s contributions, for instance through a publicly available list of ‘top contributors’

3 building a safe-to-contribute environment, where the insertion of incomplete information or work-in-progress is encouraged

4 promoting team spirit.

6 Conclusions

In this paper, we propose a novel mechanism for coordinating and enhancing the collaboration results achieved by corporate communities, in the field of tacit knowledge content production. The proposed mechanism uses the CI of the corporate community to coordinate the latter in a swarm-resembling manner, distributing and allocating the knowledge resources of its individual members, in such a way as to achieve more efficient knowledge production.

Experimental results, obtained through organisational simulation indicate that the proposed coordination mechanism can help the corporate community achieve better results in terms of content quality and timeliness, compared to the respective results achievable based only on the fully self-coordination pattern of current collaborative content creation systems, like corporate wikis.

Future work includes examining more organisational scenarios and a variety of constraints, organisational policies and resource allocation schemas.

Acknowledgements

This work was partially carried out during the tenure of an ERCIM 'Alain Bensoussan' Fellowship. The present project is supported by the National Research Fund, Luxembourg.

References

Arrow, K.J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J.O., Levmore, S., Litan, R., Milgrom, P., Nelson, F.D., Neumann, G.R., Ottaviani, M., Schelling, T.C., Shiller, R.J., Smith, V.L., Snowberg, E., Sunstein, C.R., Tetlock, P.C., Tetlock, P.E., Varian, H.R., Wolfers, J. and Zitzewitz, E. (2008) 'The promise of prediction markets', *Science*, Vol. 320, No. 5878, pp.877–878, DOI 10.1126/science.1157679.

Barney, J.B. (2001) 'Resource-based theories of competitive advantage: a ten-year retrospective on the resource-based view', *Journal of Management*, Vol. 27, No. 6, pp.643–650, DOI 10.1177/014920630102700602.

Blankenship, S.S. and Ruona, W.E.A. (2009) 'Exploring knowledge sharing in social structures: potential contributions to an overall knowledge management strategy', *Advances in Developing Human Resources*, Vol. 11, No. 3, pp.290–306, DOI 10.1177/1523422309338578.

Cao, L., Luo, D. and Zhang, C. (2009) 'Ubiquitous intelligence in agent mining', *Agents and Data Mining Interaction: 4th International Workshop, ADMI 2009*, Budapest, Hungary, 10–15 May 2009, revised selected papers, Springer-Verlag, pp.23–35.

Chesbrough, H.W. (2003) *Open Innovation: The New Imperative for Creating and Profiting from Technology*, Harvard Business School Press, Boston, USA.

Corso, M., Martini, A., Paolucci, E. and Pellegrini, L. (2003) 'Knowledge management configurations in Italian small-to-medium enterprises', *Integrated Manufacturing Systems*, Vol. 14, No. 1, pp.46–56.

Cosley, D., Frankowski, D., Terveen, L. and Riedl, J. (2007) 'SuggestBot: using intelligent task routing to help people find work in wikipedia', *Proc. Proceedings of the 12th International Conference on Intelligent User Interfaces*, ACM.

Curado, C. (2006) *The Knowledge Based-view of the Firm: From Theoretical Origins to Future Implications*, ISEG – Department of Management, Universidade Técnica de Lisboa, Lisboa, Portugal.

Engelbrecht, A.P. (2005) *Fundamentals of Computational Swarm Intelligence*, John Wiley & Sons, Chichester, UK.

Figueiredo, A., Nabuco, O., Al-Chueyr, T. and Rodrigues, M. (2009) 'Framework proposal to evaluate trustworthiness in an online community', *Proc. Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, Vol. 3, IEEE Computer Society.

Greenwood, R. and Empson, L. (2003) 'The professional partnership: relic or exemplary form of governance?', *Organization Studies*, Vol. 24, No. 6, pp.909–933, DOI 10.1177/0170840603024006005.

Hao, P-Y., Chiang, J-H. and Tu, Y-K. (2007) 'Hierarchically SVM classification based on support vector clustering method and its application to document categorization', *Expert Systems with Applications*, Vol. 33, No. 3, pp.627–635.

Hasan, H. and Pfaff, C.C. (2006) 'The wiki: an environment to revolutionise employees' interaction with corporate knowledge', *Proc. Proceeding of the 20th Conference of the Computer-Human Interaction Special Interest Group (CHISIG) of Australia on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, ACM, pp.377–380.

Heckscher, C. and Adler, P.S. (2007) *The Firm as a Collaborative Community: Reconstructing Trust in the Knowledge Economy*, Oxford University Press, New York, USA.

Helbing, D., Yu, W. and Rauhut, H. (2011) 'Self-organization and emergence in social systems: modeling the coevolution of social environments and cooperative behavior', *The Journal of Mathematical Sociology*, Vol. 35, Nos. 1–3, pp.177–208, DOI 10.1080/0022250X.2010.532258.

Hernández, H. and Blum, C. (2009) 'Ant colony optimization for multicasting in static wireless adhoc networks', *Swarm Intelligence*, Vol. 3, No. 2, pp.125–148.

Hester, A.J. (2010) 'Increasing collaborative knowledge management in your organization: characteristics of wiki technology and wiki users', *Proc. Proceedings of the 2010 Special Interest Group on Management Information System's 48th Annual Conference on Computer Personnel Research on Computer Personnel Research*, ACM, Vancouver, BC, Canada.

Holland, J.H. (2006) 'Studying complex adaptive systems', *Journal of Systems Science and Complexity*, Vol. 19, No. 1, pp.1–8.

Introne, J., Laubacher, R., Olson, G. and Malone, T. (2011) *The Climate CoLab: Large Scale Model-based Collaborative Planning*, MIT Center for Collective Intelligence, Cambridge, MA, USA.

Javanmardi, S., Ganjisaffar, Y., Lopes, C. and Baldi, P. (2010) 'Statistical measure of the effectiveness of the open editing model of Wikipedia', *4th Int'l AAAI Conference on Weblogs and Social Media*, 23–26 May, George Washington University, Washington, DC.

Kang, S.C., Morris, S.S. and Snell, S.A. (2007) 'Relational archetypes, organizational learning, and value creation: extending the human resource architecture', *Academy of Management Review*, Vol. 32, No. 1, pp.236–256.

Keen, A. (2007) *The Cult of the Amateur: How Today's Internet is Killing Our Culture*, Doubleday Business, New York, USA.

Kittur, A. and Kraut, R.E. (2008) 'Harnessing the wisdom of crowds in Wikipedia: quality through coordination', *Proc. Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, ACM.

Kittur, A., Lee, B. and Kraut, R.E. (2009) 'Coordination in collective intelligence: the role of team structure and task interdependence', *Proc. Proceedings of the 27th International Conference on Human Factors in Computing Systems*, ACM.

Klein, M., Faratin, P., Sayama, H. and Bar-Yam, Y. (2007) 'Negotiating complex contracts', CCI Working Paper 2007-002, MIT Sloan School of Management Working Paper 4196-01.

Kong, N., Hanrahan, B., Weksteen, T., Convertino, G. and Chi, E.H. (2011) 'VisualWikiCurator: human and machine intelligence for organizing wiki content', *Proc. Proceedings of the 16th International Conference on Intelligent User Interfaces*, ACM.

Levy, P. (1997) *Collective Intelligence: Mankind's Emerging World in Cyberspace*, Perseus Books, Cambridge, MA, USA.

Lomi, A. and Pattison, P. (2004) 'Introduction to the CMOT special issue on mathematical representations and models for the analysis of social networks within and between organizations', *Computational and Mathematical Organization Theory*, Vol. 10, No. 1, pp.5–15, DOI 10.1023/B:CMOT.0000032626.00525.7b.

Lykourentzou, I., Dagka, F., Papadaki, K., Lepouras, G. and Vassilakis, C. (2012) 'Wikis in enterprise settings: a survey', *Enterprise Information Systems*, Vol. 6, No. 1, pp.1–53, Taylor & Francis.

Lykourentzou, I., Djaghloul, Y., Papadaki, K., Dagka, F. and Latour, T. (2011a) 'Planning for a successful corporate wiki', *The International Conference on Digital Enterprise and Information Systems, DEIS 2011*, 20–22 July, London, UK.

Lykourentzou, I., Vergados, D.J. and Kapetanios, E. (2011b) 'Collective intelligence systems: classification and modeling', *Journal of Emerging Technologies in Web Intelligence*, Special issue on 'digital ecosystems', Vol. 3, No. 3, pp.217–226.

Lykourentzou, I., Papadaki, K., Vergados, D.J., Polemi, D. and Loumos, V. (2010) 'CorpWiki: a self-regulating wiki to promote corporate collective intelligence through expert peer matching', *Inf. Sci.*, Vol. 180, No. 1, pp.18–38, DOI http://dx.doi.org/10.1016/j.ins.2009.08.003.

Lykourentzou, I., Vergados, D.J. and Loumos, V. (2009) 'Collective intelligence system engineering', *Proc. Proceedings of the International Conference on Management of Emergent Digital EcoSystems*, ACM.

Malone, T.W., Laubacher, R. and Dellarocas, C.N. (2009) *Harnessing Crowds: Mapping the Genome of Collective Intelligence*, MIT Center for Collective Intelligence.

Nagar, Y. and Malone, T.W. (2011) 'Combining human and machine intelligence for making predictions', Presented at *NIPS 2010 Crowdsourcing and Computational Social Science Workshop*, Whistler, Canada.

Nickerson, J.A. and Zenger, T.R. (2004) 'A knowledge-based theory of the firm – the problem-solving perspective', *Organization Science*, Vol. 15, No. 6, pp.617–632, DOI http://dx.doi.org/10.1287/orsc.1040.0093.

Onnela, J-P. and Reed-Tsochase, F. (2011) 'Spontaneous emergence of social influence in online systems', *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 107, No. 43, pp.18375–18380 (published online before print), DOI 10.1073/pnas.0914572107.

Ortega, F. and Barahona, J.M.G. (2007) 'Quantitative analysis of the wikipedia community of users', *Proc. Proceedings of the 2007 International Symposium on Wikis*, ACM.

Pugh, J. and Martinoli, A. (2009) 'Distributed scalable multi-robot learning using particle swarm optimization', *Swarm Intelligence*, Vol. 3, No. 3, pp.203–222.

Rouse, W.B. and Boff, K.R. (2005) *Organizational Simulation*, p.647, John Wiley & Sons, New Jersey, USA.

Snijders, T.A.B., van de Bunt, G.G. and Steglich, C.E.G. (2010) 'Introduction to stochastic actor-based models for network dynamics', *Social Networks*, Vol. 32, No. 1, pp.44–60.

Sorensen, J.B. and Stuart, T.E. (2000) 'Aging, obsolescence, and organizational innovation', *Administrative Science Quarterly*, Vol. 45, No. 1, pp.81–112.

Surowiecki, J. (2004) *The Wisdom of the Crowds: Why the Many are Smarter that the Few*, Abacus, New York, USA.

Tan, B. and Hung, H-c. (2006) 'A knowledge management system introduction model for small- and medium-sized enterprises', *Int. J. Management and Enterprise Development*, Vol. 3, Nos. 1/2, pp.53–69.

Tang, L. and Liu, H. (2010) *Community Detection and Mining in Social Media*, p.137, Morgan & Claypool Publishers, San Rafael, CA, USA.

Tapscott, D. and Williams, A. (2008) *Wikinomics: How Mass Collaboration Changes Everything*, Atlantic Books, London, UK.

Vergados, D.J., Lykourentzou, I. and Kapetanios, E. (2010) 'A resource allocation framework for collective intelligence system engineering', *The International ACM Conference on Management of Emergent Digital EcoSystems, MEDES 2010*, Bangkok, Thailand, 26–29 October.

Wagner, C. (2006) 'Breaking the knowledge acquisition bottleneck through conversational knowledge management', *Information Resources Management Journal*, Vol. 19, No. 1, pp.70–83.

Wagner, C. and Schroeder, A. (2010) 'Capabilities and roles of enterprise wikis in organizational communication', *Technical Communication*, Vol. 57, No. 1, pp.68–89.

Wooldridge, M. (2009) *An Introduction to Multiagent Systems*, John Wiley & Sons, Glasgow, UK.

Yates, D., Wagner, C. and Majchrzak, A. (2010) 'Factors affecting shapers of organizational wikis', *Journal of the American Society for Information Science and Technology*, Vol. 61, No. 3, pp.543–554, DOI 10.1002/asi.21266.

Zacharias, G.L., MacMillan, J. and Hemel, S.B.V. (2008) *Behavioral Modeling and Simulation: From Individuals to Societies*, National Academies Press, Washington DC, USA.

Zhang, Y., Sun, A., Datta, A., Chang, K. and Lim, E-P. (2010) 'Do Wikipedians follow domain experts?: a domain-specific study on Wikipedia knowledge building', *Proc. Proceedings of the 10th Annual Joint Conference on Digital Libraries*, ACM.