Social Science Computer Review

http://ssc.sagepub.com

Evaluating Behavioral Change in Multigroup Collaboration for Content Publishing Over the Web

Charalampos Z. Patrikakis, Maria Koukouli, George K. Papadopoulos and Alexander B. Sideridis

Social Science Computer Review 2009; 27; 59 originally published online Jun 5, 2008;
DOI: 10.1177/0894439308319449

The online version of this article can be found at:
http://ssc.sagepub.com/cgi/content/abstract/27/1/59

Published by:
SAGE
http://www.sagepublications.com

Social Science Computer Review
Volume 27 Number 1
February 2009 59-75

10.1177/0894439308319449
http://ssc.sagepub.com
hosted at
http://online.sagepub.com

Evaluating Behavioral Change in Multigroup Collaboration for Content Publishing Over the Web

Charalampos Z. Patrikakis
Maria Koukouli
George K. Papadopoulos
Alexander B. Sideridis
Agricultural University of Athens

In this article, an evaluation of the behavioral change of groups collaborating towards the production of multilingual digital content is provided by using the data publication duration times as a proxy. In parallel, the assessment of feedback as a stimulant of social behavior in developing such systems is given. The evaluation is based on a system developed by a consortium of experts in the context of a European project towards the provision of an eServices platform on Organic Agriculture (OA). Group behavior of the partners involved is studied and analyzed using regression models for the extraction of trends and patterns while the results of their collaboration are discussed and valuable conclusions regarding the improvement of collaborative work over the Internet are presented.

Keywords: *collaborative group work; social behavior; digital content; e-services; web services*

Over the last decade, the World Wide Web has been constantly expanding and is now recognized as a powerful medium offering the tools for the creation and management of multilingual and multicultural environments. The evolution of the offered services goes far beyond the basic communication, fully exploiting the Internet's potential. A great amount of information is being organized in online digital libraries, which offer a practical and comprehensive tool for the provision of access through a single point worldwide. Furthermore, the implementation of such tools leads to dissemination of knowledge beyond any limits and borders, creates cultural awareness, and contributes to the empowerment and strengthening of communities (Downie, 2003; Whiten, Loots, Trujillo, & Bainbridge, 2001).

However, in spite of the existence of several multilingual platforms which have been successfully implemented (IEEE Computer Society Digital Library Home, 2007; Oxford

Authors' Note: The ideas presented in this article have been inspired by the work performed in the context of the European eContent Bio@gro project, partially funded by the European Commission. Correspondence concerning this article should be addressed to Charalampos Z. Patrikakis, Informatics Laboratory of the Agricultural University of Athens, Iera Odos 75 Street, Greece, Athens 11855; e-mail: bpatr@telecom.ntua.gr.

Digital Library, 2007), the need for more and specialized platforms still remains as new areas of interest emerge.

Though this need for the design, implementation, and operation of such platforms still exists, there is no framework for evaluating the level of cooperation between the different groups that collaborate in the process of multilingual content publication. On the other hand, in sociology, biology, and psychology the behavioral patterns of individuals and groups working, learning, and cooperating toward a common goal or inside the same limited environment have been thoroughly analyzed (Boyd & Richerson, 1988; Cavalli-Sforza &Feldman, 1981; Giraldeau, Caraco, & Thomas, 1994; Lefebvre, 1995; Rogers, 1995). So has the information seeking behavior, which is in many cases a constituent part of the overall process of multilingual content collection and publication (Barahas & Higueras, 2003; Ellis, Cox, & Hall, 1993).

In this article, the behavioral change of groups collaborating towards the production of multilingual digital content is investigated using data publication duration time as a proxy and using as a basis the evaluation of the behavior of cooperating groups of agents, animals, and organisms as these have been studied and reported in the bibliography (Reader, 2004). The results of related work as regards information seeking behavior and cooperative work are taken under consideration for the description and conducting of experiments followed by evaluation and discussion of results over a real case study: the implementation and operation of a multilingual content collection platform for Organic Agriculture (OA; Bio@gro project Home, 2005). It should be noted that by the term platform we describe an integrated set of applications consisting of a web server, a content database, an offline content processing and metadata annotation tool, and the corresponding interfaces for users contributing, processing, and accessing the stored information. Furthermore, the reasons for selecting OA as a case study should be explained: OA is an international movement generated through the sensitization of the global community towards environmental preservation and assuring of food quality (European Commission, 2003). The selection of OA as a case study in this article is based on the fact that although OA experienced a significant development in many countries finding relevant and accurate OA information is often difficult (e.g., due to language obstacles), time-consuming, and requires access to multiple resources. Therefore, the design, implementation, and content population of a portal on OA, with multilingual, certified content of high quality, certified by OA experts is a challenging task, especially as regards cooperation of all contributing parties.

The next section of the article is dedicated to the presentation of the evaluation framework, giving a description of the methodology used, together with a presentation of the European project Bio@gro that has been used as the test bed for the needs of our work. The evaluation and the corresponding results are provided in the next section. The article concludes with recommendations on improving the performance of collaborating groups for specialized and certified content publication over the Web.

Description of the Evaluated Framework and Adopted Models

Several models have been proposed regarding information seeking behavior and have been thoroughly analyzed in the bibliography. Among them, the two most cited are that of Ellis

(Ellis, 1989; Ellis et al., 1993) and Kuhlthau's (1991). The above models are based on empirical research which has been studied in subsequent studies, which gives them an advantage compared to all other, more theoretical approaches, while they function at different levels of the overall process of information seeking (Wilson, 1999). Ellis' model, though developed for a nondigital environment, can be adapted to the on-line environment (Barahas & Higueras, 2003), although it has been developed according to the theories of cognitive psychologists (Shankar, Kumar, Natarajan, & Hedberg, 2005). However, Kuhlthau's work may be considered complementary to that of Ellis, because it attaches to the stages of information seeking, associated feeling, thoughts and actions as well as the appropriate information tasks. This latter feature is important to our evaluation study as the identification of actions and tasks associated with each level of information seeking is important for analyzing the process of information collection in distinct phases. Through this analysis, information collection can be easily mapped to corresponding steps in the overall content collection and publication process. For this, the Kuhlthau's model has been adopted. According to the model, there are six stages in the information seeking and presenting. In the case of identifying the different stages in the design and implementation of a project for the provision and support of a data collection or digital library, the Kuhlthau's model is still valid. Therefore, it has been used to map all six stages (initiation, selection, exploration, formulation, collection, and presentation) to a specific part of the information seeking process in the Bio@gro project. The description of how these six stages of the aforementioned model have been adopted in our case, are presented in Table 1.

Finally, the issue of in-group cooperation of its members has also been addressed. According to Hansen and Järvelin, human collaborative activities show a pattern that comprises of asking colleagues both internally and externally regarding experiences, and search strategies (Hansen & Järvelin, 2005). Moreover, O'Day described four levels of information sharing in group situations: sharing results with other members of a team; self-initiated broadcasting of interesting information; acting as a consultant, handling search requests made by others; and archiving potentially useful information into group repositories (O'Day &

Table 1
Description of the Content Collection Framework

Stage	Process
Initiation	Identification of the informational framework of the data to be collected, which is the area of organic agriculture (OA).
Selection	Selection of the specific electronic services and content categories regarding OA that should be supported.
Exploration	Search for information sources (not content) through libraries, the Internet, university courses and seminars, and results of research projects.
Formulation	Specialization of content collection process according to user requirements, as these have been recorded through surveys of user needs and experts' opinions.
Collection	Actual content collection through traditional and electronic sources to complete the data collection about OA to be used in the platform perspective of the topic.
Presentation	Preparation of summaries of the content, the translation and localization of the content, and the insertion in the platform in terms of data and metadata are performed.

Jeffries, 1993). In the Bio@gro project, the above levels have been adopted and translated to the necessary guidelines for in-group collaboration of group members. As a result, each group acted independently as regards the collection of information for each particular country and corresponding language, while any internal group assistance in searching, was performed within the group, or by the inclusion of extra associates, that operated as members of the information collection and processing group.

Taking into account the above, the work presented in this article is focused on the behavioral change of the groups that have collaborated in the context of the Bio@gro project for the implementation of the four latest stages of the Kuhlthau's model. The two first stages, namely the initiation and selection, have been completed before the start of the first evaluation period. In our evaluation, we have tried to translate the performance of groups, as this is recorded through the volume of published content over time, and tried to translate the findings by using proposed models for the description of social and asocial behavior.

Before we proceed with the assumptions we have made, we should first elaborate further on the past work regarding the definition of social and asocial behavior. According to Roper (1986), a sigmoid function can be used as a tool for diagnosing social transmission eliminating the possibility of asocial learning behaviors. Further research has been performed on whether distinguishing learning behavior according to the functions which describe the cumulative distribution is feasible in each case (Cloutier, Newberry, Honda, & Alldre, 2002; Galef, 1990; Rendell & Whitehead, 2001; Shettleworth, 1998). However, new studies that deal with these assertions raise doubts on this method of identifying or predicting social or asocial learning behavior (Day, Kendal, & Laland, 2001; Galef, 2001; Huffman & Hirata, 2003; Lefebvre, 1995; Reader, 2004; Wilson, 1999). Based on Reader's conclusions, distinguishing social and asocial behavior based on diffusion curves is a rather challenging procedure. Sigmoidal diffusion curves may suggest either social or asocial learning behavior and the case of using accelerating curves as a diagnostic tool involve a set of special parameters that must be taken into account (Reader, 2004). However, the fact in all the above research work is that accelerating curves, such as the sigmoid, represent learning behaviors.

The aforementioned results have been taken into consideration in the evaluation of the behavior of groups and the study of behavioral changes, as these were measured over the change or improvement of publication times. Conclusions about the social or asocial behavior of these groups were drawn on this basis.

The Bio@gro eServices Platform

The Bio@gro platform has been the product of an European Commission funded project, aiming at the development of an OA eServices System (Sideridis, Costopoulou, Patrikakis, Manouselis, & Stalides, 2005) in the framework of the eContent programme that aimed at making digital content in Europe more accessible, usable, and exploitable. In this context, the Bio@gro project has developed an e-services platform, consisting of several modules, the most important of which has been the Information Catalogue (Patrikakis, Sideridis, Konstantas, & Stalides, 2006), for implementing a data collection covering several topics related to OA. The majority of this collection constituted information that was already published over the Internet, following the requirement of the eContent framework for the provision of new platforms that facilitate access to existing public and certified information, rather than the creation of new content. Each entry of information called BCO (Bio@gro Content Object) is available in the platform in four languages (English, Greek, German, and Romanian) and covers the following categories: news and events, legislation and certification related information, scientific articles, studies and best practice guides, directory of OA organizations and agencies, farmers, suppliers and processors, market reports, educational material, and online courses. For the collection of content, a set of quality criteria were followed regarding the relevancy to OA, the reliability (of source and content), regional relevancy, and timeliness (Lang, Eppler, Sideridis, & Patrikakis, 2005).

To ensure an effective work flow management during the content collection process, different groups of users were created. A group of experts (Bio@gro Editors and Publishers) in organic farming was created to manage the collection of content, while another group (Steering Committee) of highly experienced academics and professionals was assigned to peer review and evaluate the published content. The users of the system were agronomists or researchers in the field of Agriculture. Therefore, lack of knowledge and familiarity in using Internet technologies has been be taken into account as this has been recorded in the user requirements capturing phase (Patrikakis et al., 2006, Patrikakis, Sideridis, Koukouli, Manouselis, & Kostopoulou, 2006). Furthermore, the users had to be introduced with the term of metadata and also to be familiarized and trained to insert the collected information according to the metadata schema that was adopted (Dublin Core Application Profile [DCAP] Guidelines, 2005). For this reason, all the members of the above groups were trained in collecting the metadata for the BCOs, the use of the online mechanism, translation requirements and norms, before the initial phase of the content collection and evaluation process. Therefore, the collaboration between the groups of users was imperative and the way this collaboration was performed is a part of this study. The result has been a new platform that, though it had been based on existing software, was in fact a completely new system incorporating new modules, methods for data handling, new user interfaces, and work flow support mechanisms.

For tracking and collecting the content, two different groups (Bio@gro Collectors & Modifiers) were created from each participating country and were assigned with the task of providing (and translating/transforming) specific content, either from their own specific field of work (e.g., horticulture, poultry, market reports) or from other content categories. All the above roles are described graphically in Figure 1, where the relationships between all entities are highlighted. In the figure, the different actions and responsibilities of each actor group, as well as the interfaces between them are depicted, following the work flow of information processing from the initial collection to the final publication and external auditing.

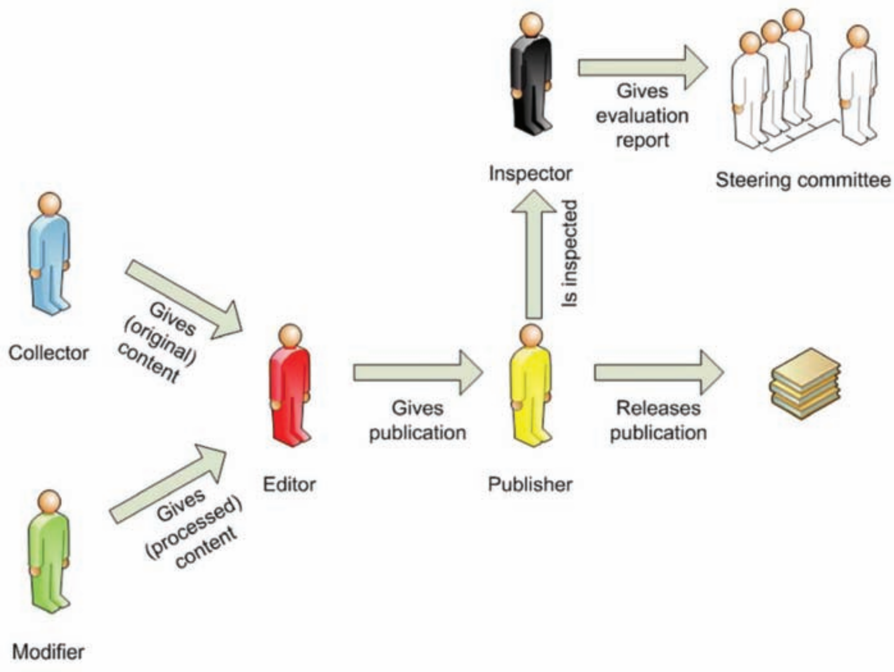
The Bio@gro platform under study here can be characterized as a Distributed Collective Practice (DCP) rather than simply a Computer Supported Cooperative Work (CSCW). The reason for this characterization is that the term DCP is used to signify collective activity mediated through geographical and conceptual distances, time, collective resources, and heterogeneous perspectives or experiences (Turner, Bowker, Gasser, & Zacklad, 2006), covering a broader area of activities than CSCW systems, which are focusing on practice and not products (Schmidt, 2000). As regards the issue of collaboration between heterogeneous groups in the context of a DCP, the subgoals of heterogeneous actors may vary based on their own understandings of the overall goal. Therefore, the actual work of the collective practice can happen only when there are immediate alignments of individual or organizational objectives with the larger arena of practice (Cragin & Shankar, 2006). Starting from the conclusions of Cragin and Shankar, we can see the importance of evaluating the behavioral change (and in parallel the changes in the overall performance) of actors participating in a collaborative practice and the efficiency of all groups with respect to the overall objective.

Content Collection and Publication Process

According to the methodology used, the design of the content collection, evaluation, and publication process was separated into two phases: the first allowed groups to work independently, with the minimum level of communication, not including exchange of knowledge or information regarding performance and good or best practices, whereas the second incorporated the necessary feedback mechanisms for indicating problems in performance and quality, as well as inadequacies of some groups compared to others. In both stages, the work performed by the participating groups was identified as the one described by the four latter stages of the Kuhlthau's model, namely, exploration, formulation, collection, and presentation (Kuhlthau, 1991). The inclusion of the feedback mechanism has been performed in the context of a process that tried to integrate at the second phase a collective conscience to the system, similar to that encountered in social groups, where collective knowledge is used to improve performance in specific tasks through social learning (Reader, 2004). As Wilson (1999) very well points in his study on models in information behavior research: "information processing and use is shown to be a necessary part of the feedback loop, if information needs to be satisfied" (p. 256). For this, the prime concern when designing the feedback mechanism was the inclusion of information about the output results that covers both the expert's opinion on the produced content as well as the user's feedback regarding access and use of it. Therefore, the feedback mechanism can also be considered as a mean of diffusing best practice experience among the groups. By the term "best practices," we refer to a set of guidelines, together with the corresponding application toolset used to assist the users in data entry of the collected information, such as the metadata to be provided for each object and the specific rules for completing the corresponding fields in the forms used in this procedure. The experience has shown that even if the metadata schema for the content collection is specific and clear, reaching a point where all users have a common understanding in completing the metadata is a difficult task. This is due to many reasons such as inexperience with the use of metadata provision and processing tools, different approaches and ideas on the same issues, and lack of a common marking or evaluation scheme (what is considered very good for me may be simply average to another). Therefore, the provision of additional assistance in the form of guides and examples is imperative.

In the case presented in the article, best practices guidelines were created by the cross examination of the reports collected as well as the internal communication of users when special issues were raised regarding faulty entries that had to be removed and republished after modification. The use of alerts during the second phase offered the tool for identifying different behaviors among the users as regards their response to an alert as well as the quality of their work.

Figure 1
The Specific Roles of Actors With Respect to the Publication Procedure



Evaluation Tests and Discussion of Results

Setting up the Trials and Defining the Evaluation Toolset

The evaluation phase was divided in two periods corresponding to the deployment of the two phases described earlier in the article. In the first period, only the first phase of the content collection, evaluation, translation, and publication mechanism was deployed. After a BCO was published, no feedback to the system was provided. In the second, the stimulating mechanism was used, based on feedback from the postpublication evaluation of the content. The first period started on July 27, 2006 and ended on October 17, 2006, while the second period started on November 16, 2006 and ended on December 22, 2006. During the first period, content collection, editing, and publication was performed in a pipelined mode.

For both periods, a reporting mechanism was installed that recorded the time needed for the publication of a BCO from the moment it was initially introduced in the platform until the time it was published. Hence, for each BCO in the system, a record containing the action performed (introduction to the system, modification, publication), the BCO's unique ID in the system (referred to as X in diagrams), the period passed until its publication in days (referred to as T in the rest of the article), and the ID of the person responsible for this action were recorded. In detail, the report notifications consisted of three different message types, as follows: A “NEW” report notification message, sent whenever a new BCO was inserted in the database. An “UPDATE” report notification message, sent whenever a BCO was modified (e.g., translated) by a user. Finally, a “PUBLISH” report notification message, sent whenever a BCO was certified and finally published in the platform. This information was then analyzed accordingly to identify the steps followed until its final publication. During the first phase, the users focused on learning the use of the system and its capabilities. For this reason, there were a number of test entries although as the results indicate the attention of the users was mainly to the familiarization with the content uploading mechanisms rather than the translation of content from other users of the system. Furthermore, at the second period a specific goal of publishing a concrete number of BCOs until the end of the project was set to monitor and evaluate the performance of the collaborating groups with respect to specific targets (number of BCOs published). To support this, monitoring of the number of BCOs published in the system was used, together with an alerting mechanism to stimulate and intensify the work flow whenever this was necessary. The mechanism was based on the provision of reports on the work done on a regular basis and the dispatching of alerts to groups that lacked in performance. Through these reports, users were also informed about necessary modifications in a way that a set of guidelines for best practices was initiated through this procedure. This set of guidelines in combination with the internal communication in each group had set up the collaborative environment that resulted in a more tuned work flow during this phase.

The analysis of the data collected for the two evaluation periods was based on the number of publications produced over a specific time period. The report notifications for the BCOs inserted in the platform were used for this analysis. The goal was to try and determine the behavior of the system in terms of performance measured over the time needed for a BCO to be published. The data collected for the two periods were analyzed and regression models for both periods with a R^2 (coefficient of determination) were produced. The statistical analysis of the collected data to find a stochastic model for each subperiod and the related prediction models was performed by using the statistical package StatGraphics Plus v.4.0 applying a Simple Regression model for the 1st period and a Linear Regression model with indicator variables for the 2nd period. The results, together with the corresponding regression models for each case are given later in the article.

A Comparative Presentation of Results for the Two Periods

Processing of the log files for the notifications sent regarding the actions for content upload and publication, reveals that in the first period there have been 136 new uploads, 529 updates, and 45 publications. The corresponding numbers for the second period are 375 uploads, 1,069 updates and 61 publications. The numbers indicate that the work in the second period was intensified, and that all corresponding indicators have been multiplied. Taking into account that the second evaluation period lasted much less than the first, we see that the intensification of work is much higher. As regards the increase in the number of new uploads, this can be partially justified by the fact that the contributors were more experienced on the second period, and were more acquainted with the platform. Therefore, work could proceed much faster. Furthermore, the feedback to the system provided by postpublication reports was causing resubmission of some objects, reported as NEW actions, contributing also to the increase of the number of new uploads. As regards the number of new entries, the results show that these have almost tripled at the second period. The number of updates in the second period was increased (at a lower rate than that of new uploads because some objects were modified before they were translated in all languages).

If we take a look into the system's behavior, we see that apart from an improvement that was caused by the increasing expertise regarding the platform use and the experience related to the tasks performed the results of the two periods, as regards uploading and updating new publications, present no difference and follow a linear pattern as depicted in Figures 2 and 3. However, when examining the number of publications performed and the mean time between upload and publication of a new object in the platform things are different. In both cases, the linear pattern observed in the new uploads and updates of BCOs does not appear anymore.

Furthermore, if we try to examine the average number of days between uploading an object and publishing it in the platform by examining the reports for each NEW and PUBLISH entry for each BCO collected in the reports, we see that as Figure 4 clearly indicates, the average number of days between uploading and publishing a BCO is increasing in the first period until it reaches a steady value of 67 days. At the second period, the number of days starts from this value and is decreasing until it is stabilized at the value of 39 days. To identify the behavior of the collaborative groups in the two evaluation periods, we have examined the time it took for specific BCOs from their first appearance to the platform as new entries to their final publication. Regression models were subsequently used to represent the patterns observed. In the following paragraphs, the results of this work for the two periods are presented.

1st Period Analysis

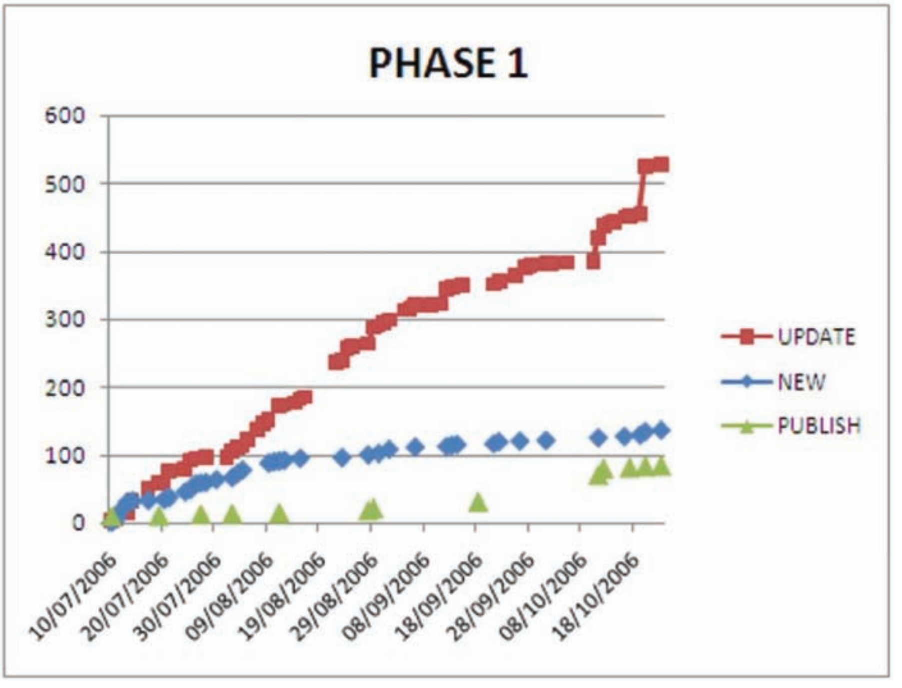
During this period, a number of 45 BCOs published within 83 days was examined. Figure 5 depicts the time needed for each BCO from the time it was inserted in the system until it was published, during the first period of the evaluation phase. The horizontal axis represents the IDs of BCOs inserted in the system in chronological order (number X corresponds to the Xth BCO inserted in the system), whereas the vertical axis represents the time interval between the insertion of the BCO in the system in its original language until the final publication in the multilingual form. At the same figure, the results of fitting a regression model to describe the relationship between the system and the publication time are also displayed by a continuous line. The fitted model is an S-curve with $R^2 = 95.80\%$ described by the following equation:

$$T = e^{\left(4.5 - 4.8\frac{1}{X}\right)} \tag{1}$$

Trying to identify the behavior of the overall system, we have reverted to the global bibliography regarding the explanation of behavior of groups of agents, humans, and animals which is presented in the section of the analysis of literature used.

Having in mind that the first period was characterized by the independent work of all groups, we can assume that this is the result of asocial behavior. Hence, the findings of the analysis of the first evaluation period indicate that such behavior indeed exists. The S-curve

Figure 2
Phase 1-Number of Days Between Uploading an Object and Publishing it

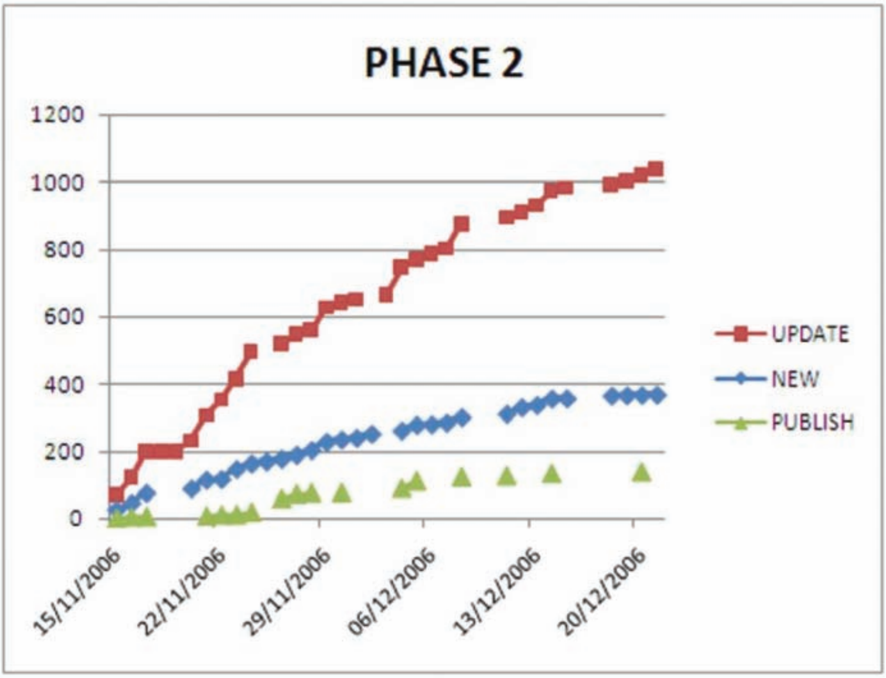


which describes with high accuracy the behavior of the system as regards the number of BCOs published over time describes a natural process (Modis, 1996). The time needed for a BCO to be published has gradually started to increase as more BCOs were inserted in the system, which is reaching saturation point. This is because the critical mass of BCOs inserted is greater than the corresponding total number of BCOs that could be processed by the system within its capacity. The next period, in which the feedback mechanism acts as the means for sharing, knowledge, and expertise among the collaborative groups, would reveal if the behavior would change and improve.

2nd Period Analysis

During second period, 59 BCOs published within 31 days were examined. Figure 6 depicts the time needed for each BCO from the time it is inserted in the system until it is published during the second period of the evaluation phase. As mentioned earlier in the article, during the second phase a stimulating mechanism was used for delivering specific alerts to the users. A time plan for the use of this alert mechanism was decided and for this reason the results of the reports received by the system were analyzed for providing the necessary feedback through the scheduled alerts. The time for sending each alert was decided according to the specific circumstances of the corresponding period (e.g., holidays, absence

Figure 3
Phase 2-Number of Days Between Uploading an Object and Publishing it

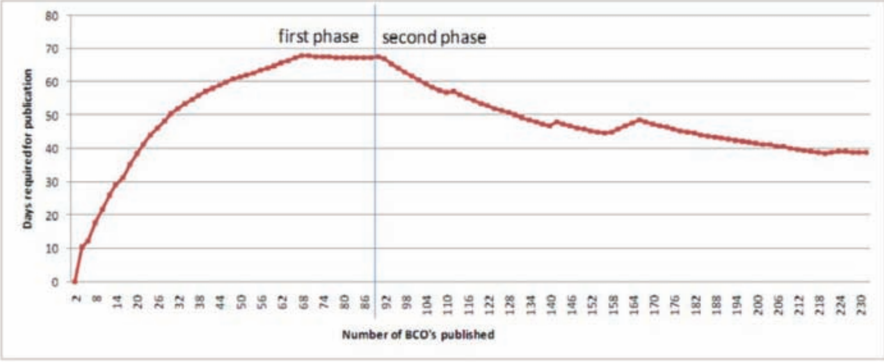


of users, special conditions, etc.). These periods between each alert were found to be reflected in the results taken, which depict the intensification of the users' effort after receiving a new alert. The data for the second period analysis are divided in six subperiods. The details for these subperiods are included in Table 2. It should be noted that the data collected for the 1st subperiod (of the second evaluation period), have been omitted from the statistical analysis that led to the production of the model which describes the behavior of the system. For this reason, equation (2) presented later includes periods 2 to 6. The reason for omitting the data of the 1st subperiod is that during the transitional phase between the two evaluation periods several BCOs that were produced during the first evaluation period remained in the system. Therefore, the second evaluation period started with some data already in the system, which resulted in deviations on the initial reports for this period. For this, the related reports were not taken under consideration.

Trying to identify the behavioral pattern for this period, the Linear Regression model with indicator variables was applied having $R^2 = 85.63\%$ and $R^2_{adj} = 82.39\%$ (adjusted R^2). The results show that the slopes of the subperiods analyzed are statistically different ($P < .0001$). The equation of the fitted model is:

$$T = -15.77 + 1.59X - 1.84I_3 - 39.27I_4 - 85.16I_5 - 171.3I_6 - 0.53XI_3 + 0.1XI_4 + 0.87XI_5 + 2.06XI_6$$

Figure 4
Average Number of Days Between Uploading an Object and Publishing it (Both Periods)



where X is the predictor variable and I_i, $i = 3 \ldots 6$, indicator variables, where $I_i = 1$ if the data belong to period i and $I_i = 0$ otherwise. This equation describes the following five linear models:

$$
\begin{array}{lll}
T = -15.77 + 1.59X & \text{for} & 10 \le X \le 17 \\
T = -17.6 + 1.06X & \text{for} & 18 \le X \le 30 \\
T = -55.04 + 1.69X & \text{for} & 31 \le X \le 42 \\
T = -100.93 + 2.46X & \text{for} & 43 \le X \le 49 \\
T = -187.07 + 3.65X & \text{for} & 50 \le X \le 59
\end{array}
$$

Looking at the history of the project and the log files, we found out that the passing from one period to the next one coincides with the processing of a feedback report, and the corresponding meetings with the groups related to this report. Since it is obvious that the processing of a feedback report clearly signifies the end of a period, and the start of another in which the processing and publication times are improved, the next step was to determine if the behavior of all the users is the same in all the subperiods. If the regression lines were parallel, this would indicate that the system could not be affected by external factors or special conditions. The fact that there is difference between the lines shows that the system is influenced by special conditions occurred in certain periods which influence the behavior of the users. These can be grouped in two different subgroups according to their slope.

The 2nd and 4th subperiods were further analyzed and the results show that these have no statistically significant difference in their slopes ($P = .7593$). The analysis indicates that there was no significant change in the behavior of the system for the 2nd and 4th subperiods as regards publication times. On the contrary, the 3rd, 5th, and 6th subperiods present significant difference in their slopes ($P = .0001$). Looking back at the time where this change has occurred, it was discovered that these three sub periods present a different behavioral pattern of the users because of the different circumstances and special conditions during each one.

Figure 5
Regression Model for the First Evaluation Period

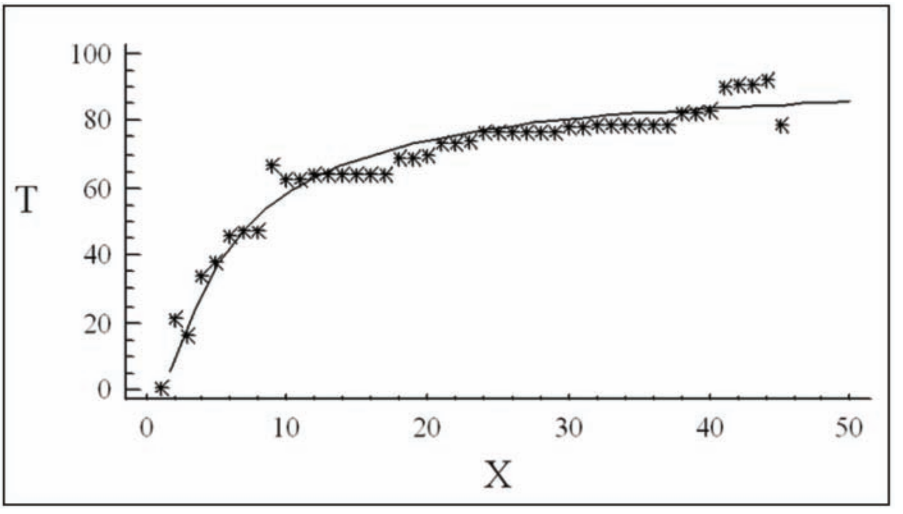


Table 2
Subperiods of the Second Period and Number of BCOs Published

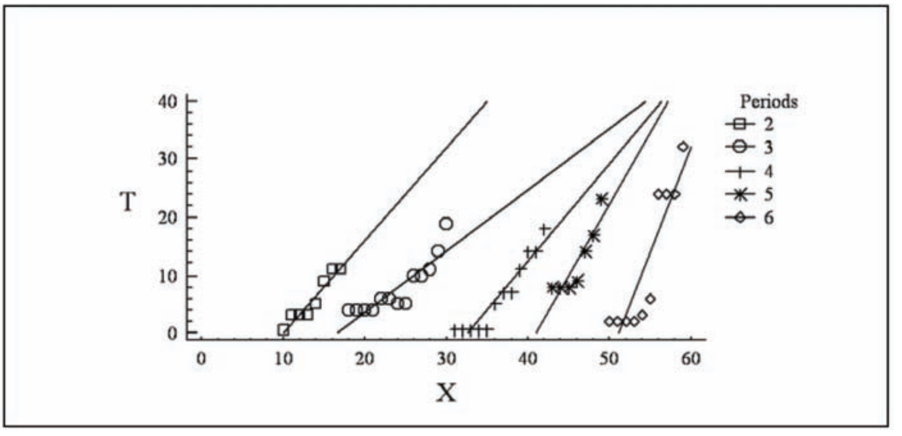
Subperiod	1st	2nd	3rd	4th	5th	6th
BCOs published	$3 \leq X \leq 9$	$10 \leq X \leq 17$	$18 \leq X \leq 30$	$31 \leq X \leq 42$	$43 \leq X \leq 49$	$50 \leq X \leq 59$

BCO = Bio@gro Content Object.

More specifically, the delay observed in the content work flow during the 3rd period is due to the users focusing their effort for the preparation of other deliverables of the project before a consortium meeting. As regards the 5th period, this refers to BCOs that were published by a single user in only one day, during which the rest users were unable to contribute because of a consortium meeting. Finally, regarding the 6th period, this refers to the time when the project was about to end and for this reason the effort of all the participants was focused on completing all the necessary deliverables and reports of the project and not only in the operation of the publication mechanism.

Therefore, concentrating on the set of the 2nd and 4th subperiod, the outcome shows that the deployment of the feedback mechanism has improved the performance of the system. In this view, it could be claimed that the feedback reports and the corresponding meetings and discussions of results have indeed brought the desired result: the good practices of some groups and the bad practices of others have been known, and subsequently, the performance of all groups was improved. As Figure 6 clearly indicates, the improvement on the system's behavior from the use of the feedback mechanism is evident. However, it should be noted

Figure 6
Regression Model With Indicator Variables for the Second Evaluation Period



that all subperiods demonstrate a pattern that is similar to that of the first period: the number of days for publishing a BCO increases gradually until the next feedback, in which a sudden decrease is encountered. Another interesting finding is that though there is no statistically significant difference between the 2nd and 4th subperiod, as the number of BCOs increases (corresponding to later time periods), so do the days needed for publishing a BCO. This indicates a relaxing of the effect that the feedback mechanism introduces and could be explained by the fact that it is gradually starting to be considered as the normal routine by all actors involved in the mechanism.

Summary and Conclusions

Undoubtedly, there is a definite need for more and varied studies of how people work together to understand the relationship between collective activity, systems implementation, and the impacts of knowledge sharing on work flows. In this article, we have studied the behavioral changes in multigroup collaboration for content publishing over the Internet through the evaluation of publication data times for an actual system. The changes in content upload and publication times for multilingual content over a web platform implemented in the context of the EU project Bio@gro were used as input to identify the behavioral pattern of the users in two different periods. In the first one, which was characterized by the independent work of all groups, user behavior was identified as an asocial behavioral process, whereas in the second one, the application of a feedback mechanism improved the data publication times. This was achieved through the triggering of a new process cycle upon the incorporation of each feedback result. The given explanation is that the incorporation of each feedback report disseminates the knowledge of best practices applied by certain user groups, together with the expertise from bad practices, so that improvement of the work performed and corrections of mistakes can take place. The above indicate that the collaboration of different groups can be improved when mechanisms for stimulating social behavior, such as periodic feedback reports and/or discussions among groups regarding practices are used. However, even the effect of these social behavior stimulating mechanisms is reduced over time, because they are gradually considered part of the work routine and their stimulation effect is minimized.

Based on the above conclusions, a list of actions toward the increasing of effectiveness of collaborative groups through social behavior stimulating mechanisms can be deployed:

- Incorporate a feedback mechanism that can inject to the system reports incorporating knowledge and expertise as regards best and bad practices.
- Monitor the behavior and overall performance of collaborative groups and whenever signs of reduced productivity rates appear, provide feedback and discussions that will trigger a new process cycle.
- Try to eliminate the routine effect that may appear in the publication process.
- Try to identify deviations from the standards behavior that may be the result of hasty actions (e.g., massive publications that may lack in quality) or even the result of an innovative approach.

The above actions may be achieved through different supporting automatic and manual mechanisms such as the use of reminder e-mails containing progress and delay reports, use of SMS messages (Virtanen, Sirkiä, & Jokiranta, 2007), and direct communication with the responsible editors. However, the danger of the supporting mechanisms becoming part of the routine is also present, especially in the case when these mechanisms are automated. For example, if the mechanism sends a reminder e-mail at the beginning of each month then the user may get used to this and consider it as just another part of his "monthly" routine. The parameter of unpredictability in the generation of reminders should be seriously taken under consideration. Finally, the identification of quantifiable performance indicators and the parallel deployment of a monitoring mechanism are imperative. However, the design should correspond to the overall data work flow for collection, evaluation, and publication.

References

Barahas, M., & Higueras, E. (2003). *SEEKS project public deliverables—Deliverable 2 "Initial Taxonomy of ISB."* Retrieved August 5, 2007, from SEEKS project Learners Information Seeking Strategies in the Information Society: http://www.ub.es/euelearning/seeks/

Bio@gro project Home. (2005). Retrieved September 10, 2007, from Bio@gro project-All about Organic Agriculture: http://bioagro.aua.gr

Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Cavalli-Sforza, L., & Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton, NJ: Princeton University Press.

Cloutier, S., Newberry, R. C., Honda, K., & Alldre, J. R. (2002). Cannibalistic behavior spread by social learning. *Animal Behavior, 63*, 1153-1162.

Cragin, M. H., & Shankar, K. (2006). Scientific data collections and distributed collective practice. *Computer Supported Cooperative Work, 15*, 185-204.

Day, R. L., Kendal, J. R., & Laland, K. N. (2001). Validating cultural transmission in cetaceans. *Behavioral and Brain Sciences, 24*, 330-331.

Downie, J. (2003). Realization of four important principles in cross-cultural digital—Workshop Paper for JCDL. *NSF Workshop on Research Directions for Digital Libraries*. Massachusetts.

Dublin Core Application Profile (DCAP) Guidelines. (2005). Retrieved December 15, 2006, from Dublin Core Metadata Initiative: http://dublincore.org/documents/library-application-profile/

Ellis, D. (1989). A behavioral approach to information retrieval system design. *Journal of Documentation, 45*, 171-212.

Ellis, D., Cox, D., & Hall, K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation, 49*, 356-369.

European Commission. (2003, June 26). *CAP reform-agriculture and rural development-EUROPA*. Retrieved September 15, 2004, from Agriculture and Rural Development-EUROPA: http://europa.eu.int/comm/agriculture/capreform/index_en.htm

Galef, B. G. (1990). Tradition in animals: Field observations and laboratory analyses. In M. Bekoff & D. Jamieson (Eds.), *Interpretation and explantation in the study of behavior: Vol 1 Interpretation, intentionality, and communication* (pp. 74-95). Boulder, CO: Westview.

Galef, B. G. (2001). Where's the beef? Evidence of culture, imitation,and teaching in cetaceans? *Behavioral and Brain Sciences, 24*, 335.

Giraldeau, L.-A., Caraco, T., & Thomas, J. V. (1994). Social foraging: Individual learning and cultural transmission of innovations. *Behavioral Ecology, 5*(1), 35-43.

Hansen, P., & Järvelin, K. (2005). Collaborative information retrieval in an information-intensive domain. *Information Processing & Management: An International Journal, 41*, 1101-1119.

Huffman, M. A., & Hirata, S. (2003). Biological and ecological foundations of primate behavioral tradition. In M. Fragaszy & S. Perry (Eds.), *The biology of traditions: Models and evidence* (pp. 267-296). Cambridge, UK: Cambridge University Press.

IEEE Computer Society Digital Library Home. (2007). Available July 20, 2007, from IEEE Computer Society Digital Library: http://www.computer.org/

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science, 42*, 361-371.

Lang, U., Eppler, U., Sideridis, A. B., & Patrikakis, C. Z. (2005). *eContent Project Bio@gro Deliverable D 3.2-Content Collection Methodology*. Athens: Bio@gro Consortium.

Lefebvre, L. (1995). Culturally-transmitted feeding behavior in primates: Evidence for accelerating learning rates. *Primates, 36*, 227-239.

Modis, T. (1996). *Predictions*. Irakleio, Greece: Crete Academic Publications (in Greek).

O'Day, V., & Jeffries, R. (1993). Information artisans: Patterns of result sharing by information searchers. *Proceedings on the ACM Conference on Organizational Computing Systems, COOCS '93* (pp. 98-107). Milpitas, CA: ACM Press.

Oxford Digital Library (ODL). (2007). Retrieved September 7, 2007, from Oxford Digital Library (ODL): http://www.odl.ox.ac.uk/home.htm

Patrikakis, C. Z., Sideridis, A. B., Konstantas, A., & Stalides, G. (2006). BIO@GRO: A multilingual e-services platform for the organic agriculture market. *4th World Congress on Computers in Agriculture* (pp. 151-156). Orlando, FL.

Patrikakis, C., Sideridis, A., Koukouli, M., Manouselis, N., & Kostopoulou, K. (2006). *Bio@gro Deliverable D 1.2 "User requirements analysis."* Athens: Bio@gro project consortium.

Reader, S. M. (2004). Distinguishing social and asocial learning using diffusion dynamics. *Learning & Behavior, 32*(1), 90-104.

Rendell, L., & Whitehead, H. (2001). Culture in whales and dolphins. *Behavioral and Brain Sciences, 24*, 309-324.

Rogers, E. M. (1995). *Diffusion of innovations*. New York: Free Press.

Roper, T. J. (1986). Cultural evolution of feeding behaviour in animals. *Science Progress, 70*, 571-583.

Schmidt, K. (2000). Distributed collective practice: A CSCW perspective—Invited talk. *Conference on distributed collective practices* (pp. 19-22). Paris.

Shankar, S., Kumar, M., Natarajan, U., & Hedberg, J. G. (2005). A profile of digital information literacy competencies of high school students. *The Journal of Issues in Informing Sciences and Information Technology, 2*, 355-368.

Shettleworth, S. J. (1998). *Cognition, evolution, and behavior*. Oxford: Oxford University Press.

Sideridis, A. B., Costopoulou, C. I., Patrikakis, C. Z., Manouselis, N., & Stalides, G. (2005). An eServices System to support information exchange among the agricultural community. *Proceedings of the 2005 EFITA/WCCA Joint Congress on "IT in Agriculture"* (pp. 376-381). Portugal: Villa Real.

Turner, W., Bowker, G., Gasser, L., & Zacklad, M. (2006). Information infrastructures for distributed collective practices. *Computer Supported Cooperative Work, 15*, 93-110.

Virtanen, V., Sirkiä, T., & Jokiranta, V. (2007). Reducing nonresponse by SMS reminders in mail surveys. *Social Science Computer Review, 25*, 384-395.

Whiten, I. H., Loots, M., Trujillo, M., & Bainbridge, D. (2001). The promise of digital libraries in developing countries. *Communications of the ACM, 44*, 82-85.

Wilson, T. D. (1999). Models in information behavior research. *Journal of Documentation, 55*, 249-270.

Charalampos Z. Patrikakis received his Dipl.-Ing. and his PhD from the Electrical Engineering and Computer Science Department of NTUA, Greece. He is currently working in the field of informatics and computer networks research and has acted as project leader and technical coordinator of several successfully integrated projects, both in national and international levels. His main research interests are in the area of IP service and application design and implementation, P2P networking, Personal Networks, Mobile computing, and context aware services. He can be reached at bpatr@aua.gr.

Maria Koukouli received her degree in physics and master's in the environmental sciences from the Physics Department of the University of Patras. She is currently working as a research and administrative assistant in European and national projects in the Informatics Laboratory of the Agricultural University of Athens. She is a member of the Greek Physicists Association. Her main research interests are in the area of Geographical Information Systems (GIS) and Remote Sensing applications for environment and agriculture. She can be reached at mkou@aua.gr.

George K. Papadopoulos is an assistant professor in Agricultural University of Athens. He received his Diploma in Mathematics from University of Thessaloniki, MSc in OR & Computer Science from University of Athens, and PhD in Systems Reliability from University of Athens. His main research interests are in the area of Systems Reliability, Quality Control, and Web-based Teaching Environments. He can be reached at gpapadop@aua.gr.

Alexander B. Sideridis is director of the Informatics Laboratory of the Agricultural University of Athens. He earned his first degree at the University of Athens and his MSc and PhD from Brunel University. He has been the project leader of more than 30 successful national and international projects and has published more than 150 scientific articles in Management & Decision Support Information Systems, Computer networking related to local and central administration activities, Advanced computational numerical modeling, Informatics & impact of computers in society, and Agricultural Informatics. He can be reached at as@aua.gr.