Abstract

Internet based services, particularly asynchronous communication services, offer an environment suited to the rise of knowledge communities. Knowledge communities, or invisible colleges, have existing for a long time, but have not always had efficient ways of communication.

Knowledge communities are heavily dynamic and are always changing in their composition and also in each member interests. The intersection of the interests of a given number of elements creates the environment for the rising of a community. Nevertheless, this community will benefit not by the intersection but by the union of knowledge of these elements. On the other hand, multiple overlapping communities are generated and one persons set of interests includes some intersection with the interests of several knowledge communities.

These communities are rather fluid and dynamic and the old processes for scholarly communication don’t seem to fairly represent these properties. Digital Libraries projects keep appearing throughout the world. Paper printed journals and proceedings are beginning to be also distributed electronically. Online scientific journals are starting to be taken seriously by scientists and by some of the major scientific publishers.

Nevertheless, online documentation per se doesn’t address the dynamic behavior of knowledge communities. Also, the problem of the undiscovered public knowledge (as called by Don Swanson) still tends to remain unattended. Most of the times we have been using old processes to solve new problems.

In our opinion, to address these issues, an online scholarly publication should have the following goals:

1. Promote the generation of knowledge communities composed of elements coming from different scientific communities - promote interdisciplinarity assuming it is a key factor on discovering public knowledge,

2. Promote, develop and maintain the communication between elements of a knowledge community,
3. Narrow the gaps between, and integrate as much as possible, formal and informal communication.

In this paper we present our research project on online publishing of a scientific journal, we discuss some issues related with knowledge communities and scholarly communication, and we propose a set of characteristics an online scientific publication should have in order to accomplish the aforementioned goals.

1. Introduction
There has been an intense research effort around electronic publishing and scholarly communication in the last decade. From the technological questions, to the legitimacy, validation and acceptance of electronic journals by academics, from the economic model of electronic publishing to copyright issues, many questions regarding the impact of electronic journals on scholarly communication have been considered in study and experimentation.

The wide availability of Internet access, and the growing use of Internet based services, has been convincing even some skeptics that electronic scholarly journals are, or can be, not only a complement but, in many cases, a good alternative to traditional paper based journals. Online scientific journals are starting to be taken seriously by scientists and by some of the major scientific publishers.

The advantages of the electronic medium, allowing the editorial process to occur faster and with less “space constraints” (the participants in the process – authors, reviewers, editors- can be scattered around the World) than the paper based journal, alongside with the cost-effectiveness of electronic distribution and access, are being explored by academics and publishers.

But, in many cases, the new electronic journals are only “translations” of traditional paper based journals, failing to take advantage of all the possibilities that the new medium can offer. As Valauskas \[1\] points out, this is particularly true with the lack of experimentation on promoting debate and discussion through communication and interactivity between authors and readers. To take full advantage of the technology already available for scholarly electronic journals, we must consider the social aspects of scholarly communication and knowledge creation. Building on Mitchell \[2\] thoughts on digital cities, we could argue that the most crucial task before us is not producing electronically deliverable “content”, but “imagining and creating digitally mediated environments for the kinds of lives that we want to lead and the sorts of communities we will want to have”.

2. Communities and Scholarly Communication
The creation and dissemination of knowledge is a social process. While individual researchers are the producers of knowledge, their work is closely related (or build upon) the research of others with whom they share values, norms, methods and reward systems.

The role of scientific communities in knowledge creation and diffusion is a well studied subject \[3\][4][5]. It is generally accepted that scientists are members of “invisible colleges”, composed by scientists with similar interests and used to share information among them \[6]. In the engineering field, some authors prefer to use the term “knowledge community”, to refer to the group of researchers that share information and ideas \[7][8][9]. Anyway, a knowledge community (the term that we will use in this paper) or an invisible college, is a group of people who have common knowledge, goals and behavioral norms, and maintain intense social contacts and interpersonal communication.

The existence of a single knowledge community, in a given area of science or research field, fail to completely describe the complexity of the social interactions and communication patterns between its members. As Constant noticed in the aeronautical field, the community is, in fact, composed of a multilevel, overlapping hierarchy of “subcommunities” \[8].

Each of the members of the community has its own corpus of knowledge and its own interests on acquiring and/or generate (or participate on the generation of) more knowledge. A
person’s corpus of knowledge may not be all inside (and generally isn’t) the corpus of knowledge of the main community to which he/she belongs. This means that a person really has acquaintance to a set of communities, one of which is predominant, in a given time.

On the other hand, there is some knowledge that is shared by persons belonging to different communities. This common knowledge between different communities has the potential for the creation of a new knowledge community, if communication between the persons that share it can be established. This means that this community will be generated from some intersection of knowledge interests of their members, but their members will benefit from the union of their knowledge. Interdisciplinary interactions can be done - the undiscovered public knowledge (as called by Don Swanson) \footnote{10} has an opportunity to show up in these knowledge communities. Each person belongs to several knowledge communities; each knowledge community arises from the knowledge of several persons. Multiple overlapping communities are generated. If we see this by the person’s point of view, a knowledge community is generated from some intersection of knowledge from several persons. If we see this from the knowledge point of view, each person’s corpus of knowledge includes some intersection of the corpus of knowledge from several communities.

There are other characteristics of knowledge communities that should be considered. Levy and Marshall \footnote{11} characterize documents as being fixed/fluid and permanent/transient, among other things. Fixed/fluid characteristic has to do with patterns of change, or changing rate of a document in a given period. Permanent/transient has to do with the lifetime of a document \footnote{11}. By analogy to documents, we presume that communities could also be characterized as being more or less fixed or fluid and more or less permanent or transient.

Once these knowledge communities are built from the knowledge of a group of persons in a given moment, there are two variables that make them change over time:

- Community corpus of knowledge is not fixed. The rate of change may vary from one discipline to another. But as each person’s corpus of knowledge is always changing (whether because of the interaction in this community or by the interaction in other communities to which the person belongs, or by him/herself knowledge construction) the community corpus of knowledge changes too.

- Community composition is always changing. Although the personal “core” interests will probably remain the same, each person’s interests may vary over time. There are always people coming and going from sharing knowledge with the community. This makes knowledge communities to change in terms of their members’ composition.

Thus, we could say that a knowledge community is rather fluid and dynamic, in terms that it is always changing. Also, depending on the subject it can be more or less transient. For disciplines that have a very rapid progression in knowledge over time, like computer science, maybe knowledge communities are more transient. For disciplines that are more stable in knowledge, communities are maybe more permanent.

As described before, formal and informal communication between its members is the “core” of knowledge communities. The formal system of scholarly communication is closely tied to structure of those communities. Before the advent of Internet services, informal communication was reduced to physical meetings, phone calls, post-its attached to documents, annotations on documents, etc. On one hand, we used to have formal communication mechanisms, like papers, conference communications and proceedings, books, monographs, etc. On the other hand we used to have informal communication mechanisms, a part of which was derived from formal communication. People read papers, reports, books, listen to communications, etc. and then talk about them, relate them, share information (verbally or written) and acquire new knowledge. In these circumstances, informal communication mechanisms are tightly connected to formal ones.

Among others, these traditional technologies and mechanisms have two main drawbacks: (1) limits to the number of people involved and (2) require physicality most of the times for
people and/or documents. The common knowledge may exist but the community may not arise, or many potential members may not participate, because they are not aware of each other nor that they share knowledge and interests. Once that they are not aware of each other, some important interactions haven’t the proper conditions to come up and this may result potentially in slower improvement of knowledge.

With Internet services available to most universities and research institutions, informal communication mechanisms like email messages, discussion forums, Usenet newsgroups, etc, are used on a regular basis and bring a new dimension to these problems. Nevertheless, as we said before, informal communication in knowledge communities is coupled to formal communication. Thus, contents created using Internet mechanisms for informal communication, will only be effective on scholarly communication if they have connection to formal communication contents.

If we have a formal communication layer composed of formal documents for scholarly

![Diagram of document-based knowledge communities](image)

**Fig.1 - Document-based knowledge communities**
communication, knowledge communities will rise from each of the documents: composed of the people that are interested on the knowledge that document contains. Overlapping knowledge communities arise, because they are built from documents that overlap in fields of knowledge and because they are composed of members that belong also to other communities. Informal communication occurs in each community, and because it is built from a formal document, the informal communication is very much related to formal communication, as it is shown in fig. 1.

3. Online Journals for effective Scholarly Communication

It is well established the crucial role of scientific journals in formal scholarly communication. Despite some differences across disciplines (mainly between “hard” and “soft” sciences), journals are generally considered as the most important mechanism of formal scholarly communication.

Electronic journals have proliferated across the net in the past few years\[12\]. However, quantity doesn’t always imply quality - by the contrary. In a rewarding system based on the quantity of publications, online journals (ojournals) proliferate across the net. It is easy to cede to the temptation of creating a journal to have one’s own papers published \[13\]. Assuming we are talking about ojournals that have their contents quality assured by a high quality peer review process, there are still some quality related questions that have to be posed: the effective use of the electronic medium as an added value to electronic publishing and to scholarly communication.

Electronic publishing, mainly online publishing, cannot be limited to imitate paper. The electronic medium and the Internet bring a new set of potentialities to scholarly communication and to scientific journals \[14\][15]. While some ojournals are merely a transposition of paper paradigm to the Internet, there are others that really take advantage of this medium. New features, like hyperlinking, multimedia, computer programs execution, etc. are being used nowadays by some ojournals\[4\]. Other features like informal communication mechanisms associated with (or referenced by) the ojournal, as discussion forums, lists of FAQ’s, etc, keep appearing. New or improved concepts of ojournals are being experienced and studied worldwide \[15\] [16] [17] [18] [12] [19] [20] [21].

Nevertheless, most of the time only a small set of these features is used and often they seem to be used because technology allows it. What happens is that generally the technology isn’t the problem; instead, it relies on instilled habits and other social considerations drawn by centuries of use of a paper-based process. History has taught us – every time a new technology emerges, we start to use it as if it was the previous one. Slowly, we start to move from one to another, discovering new applications, applying new discoveries, delighted with the adventure of the unknown. That’s what’s happening with the web. By small steps each time, we are starting to discover how to use it for publishing, and hence, how to use it for communicating.

3.1. Goals of an Online Journal

In our opinion, to effectively improve the promotion of scholarly communication, an ojournal should have the following goals:

1. Promote the generation of knowledge communities composed of elements coming from different scientific communities - promote interdisciplinarity assuming it is a key factor on discovering public knowledge (as called by Don Swanson \[10\]);
2. Promote, develop and maintain the communication between elements of a knowledge community;
3. Narrow the gaps, and integrate as much as possible, formal and informal communication.

An ojournal shouldn’t only promote, develop and maintain scholarly communication, but it
should also promote interdisciplinarity in order to help discovering public knowledge. The electronic medium has some potentialities the paper doesn’t have, and we should take advantage of them conveniently. The third goal is almost a corollary from the first two: formal as well as informal communications, in whatever form they are, exist and are real inside any scientific community [11]; knowledge communities rely very much on informal communication mechanisms - in order to have them tied up to scholarly communication we will have to support tight links between formal and informal communication forms.

It is this ability of effective communication that provides, to a great extent, the added value of a community. And this communication, that is required to be effective, can’t be reduced to the internal environment of that community - bridges to other communities should be created; interdisciplinarity should be promoted. The community shouldn’t be closed to itself; instead, it should be open to take and to give, to share.

3.2. Characteristics for accomplishing the goals
Using the taxonomy of digital library issues proposed by Nurnberg et al. as a framework [22], we propose a set of characteristics an online journal should have in order to accomplish the goals we have established before in this paper, as it is shown in table 1.

<table>
<thead>
<tr>
<th>New online journals</th>
<th>Data</th>
<th>Metadata</th>
<th>Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperdocuments</td>
<td>Personalized and dynamic</td>
<td>Effective, fast and coherent document submission and peer review system</td>
<td></td>
</tr>
<tr>
<td>Multimedia documents</td>
<td>Table of Contents and index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple versions</td>
<td>Personalized structure and contents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer program</td>
<td>Classification of several genres of documents (including informal ones)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data sets</td>
<td>Annotations and comments that can be widely shared</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New genres of documents</td>
<td>Citation of formal document or part of it</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formal communication artifacts</td>
<td>Citation of informal document</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informal communication artifacts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic documents (with forward references, personalized and updated hyperlinks)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Processes</th>
<th>Metadata</th>
<th>New online journals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic linking with external sources</td>
<td>Personalized and dynamic</td>
<td>Hyperdocuments</td>
</tr>
<tr>
<td>Personalized searching</td>
<td>Personalized structure and contents</td>
<td>Multimedia documents</td>
</tr>
<tr>
<td>Individual work support</td>
<td>Classification of several genres of documents (including informal ones)</td>
<td>Multiple versions</td>
</tr>
<tr>
<td>Share of individual work</td>
<td>Annotations and comments that can be widely shared</td>
<td>Computer program</td>
</tr>
<tr>
<td>Group and cooperative work support</td>
<td>Citation of formal document or part of it</td>
<td>Data sets</td>
</tr>
<tr>
<td>Knowledge discovery</td>
<td>Citation of informal document</td>
<td>New genres of documents</td>
</tr>
<tr>
<td>Multidirectional communication</td>
<td></td>
<td>Formal communication artifacts</td>
</tr>
<tr>
<td>Connection between informal and formal communication</td>
<td></td>
<td>Informal communication artifacts</td>
</tr>
</tbody>
</table>

Table 1 - Characteristics of an online journal using the taxonomy presented by Nurnberg et al [22], as a framework

241
Goal 1 - Promote the generation of knowledge communities composed of elements coming from different scientific communities

Most of the times scientific journals are very much specialized - scientists demand it and need it as such. This is an advantage, but it turns out to be a disadvantage when one wants people from different scientific communities to share knowledge, when one wants interdisciplinarity. To promote interdisciplinarity, a journal should have a large scope, a large base of documents from several fields or even from several areas in order to provide common access to same contents for specialists in contiguous areas and even on apparently disclosed areas.; one the other hand, for scientists to be interested on its contents, its documents should be special purpose.

Such a journal, with so many areas, would present to the reader such a number of documents that it would be hard for him/her to choose the appropriate documents - it would be a great motivation not to use the journal. Search mechanisms could be used, but it would be a better idea to make the journal personalized based on the reader’s profile. This profile could be constructed dynamically accordingly to the user’s interaction with the system.

Goal 2 - Promote, develop and maintain the communication between elements of a knowledge community

Publishing policies should be supported by mechanisms that make it effective and efficient. We think that an ojournal would benefit from being on the web if, among other things, its contents could be updated on a very regular basis. By this we mean that a document should be published as soon as it has been accepted (afterwards, the published papers could be assembled in a formal issue for archiving and reference purposes). The journal shouldn’t have a maximum number of documents published. The electronic medium doesn’t have the space limitations that the paper has - we should take advantage of this feature.

Also, hypertext references should be maintained up to date, using mechanisms that verify and update them. Mechanisms of knowledge discovery of contents in the web for later inclusion in the journal should also be provided.

It should be easy for authors to submit documents. Presuming the journal is peer reviewed, this process should be supported by mechanisms that assure the necessary confidentiality and coherence. Editors, reviewers and authors should be provided with an easy to use and secure mechanism for efficient communication.

There are several genres of documents that could be considered for work activities of a community of knowledge workers. We should take advantage of the online potentialities and use other, eventually new, genres of documents, flexibilize our notion of document, mainly for informal communication purposes. It should be taken seriously the use of several genres of documents with different degrees of permanency and fluidity, suited for several technology constellations and work practices. Internal document structuring, allowing the use of individual article components, could provide an important added value to the overall system’s usage. Also, new documents’ genres (or part of documents) classification, such as hyperlinks, or article components, would provide new dimensions on flexibility and dynamism of the journal.

We should adapt ourselves to the fluidity the web provides us. Multi-versioning mechanisms and the inclusion of forward references (like those defined by Holoviak) for some genres of documents could be supported. Multimedia documents, data sets, computer programs and multi-language support for documents could also be provided.

In order to allow and to promote communication between community members, individual work must be fully supported. Generally, knowledge workers first work individually and afterwards they share their ideas and developments with others. Many times individual
work is the basis for the initiation of group work. In order to promote group and cooperative work, individual work practices should be seriously considered.

In order to support effective individual work, an online journal should provide mechanisms that support and make it easy to access and read documents online, offline and in paper as well as mechanisms that allow the utilization and manipulation of several file formats (also using automatic file compression/decompression). Individual personalized and general search mechanisms should be provided as well as personalized individual marking up, annotation and commentary features and statistics utilities for general use.

An ojournal would have added value if it supported group and cooperative work activities. Also the sharing of individual work within a community or part of it, should be an item to put in question.

Goal 3 - Narrow the gaps, and integrate as much as possible, formal and informal communication

In order to allow and to promote formal communication, related genres of documents should be considered (papers, book reviews, etc…). The system should allow the making of relationships between formal documents (forward references and hyperlinks to other formal related documents) and show them to the user.

To allow informal communication, it should allow informal communication document genres (comments to the journal, to the document or to part of it) and the communication between several intervenients in several directions (not only author-reader direction). An easy relationship between formal and informal communication must carry out the use of informal communication systems and structures that systematically call the usage of formal documents as a framework. Also, simple document (or document part - a section of an article, for instance) citation and reference mechanisms are needed.

4. Informattica Online

4.1. General Description

Informattica Online is a project being developed at Universidade do Minho with the collaboration of the University of Sheffield.

It works as shown in fig. 2.

The journal will be, in fact, a personalized front-end to a large database where documents will be stored: eventually each user will see a different journal accordingly to his/her profile, which includes the representation of his/her scientific interests. Documents and users’ interests will be classified using the same classification systems, allowing a match between them. Each document might be accessed by different users with potentially different profiles. This allows knowledge to be shared by persons with different profiles and promotes interdisciplinarity. A document based knowledge community is expected to rise with members that may or may not have the same profiles. Multiple overlapping knowledge communities rise from multiple documents that have overlapping areas on their classification. Each person shall belong to several of these communities. Interaction between members of a community and interaction between communities will be provided using informal communication mechanisms.

Informattica Online is being designed to have all the characteristics described in table 1, except for the Group and Cooperative work support that is previewed to be done in a later phase.
Fig. 2 - General description of Information Online

- Individual work sharing
- Discussion forums
- Communications / collaborations
- Individual work and collaboration

- Common user sees documents in journal
- Each user sees a different Dynamic Profile of contributions

Collaborations

DB3

People

DBS

Articles

DB1

Journal

Journal

Non Periodical

Personalized in Form and

Other web pages or

Database

Peer reviewed documents through fully electronic peer review system.
4.2. System’s Modules and General Functionality
The system will be divided in five main modules as shown in fig. 3.

![Fig. 3 - Main modules of the system](image)

**Submission and Peer Review**
Submission and peer review relates to all phases from when an author submits a document, until it is prepared for publication. When an author submits a document, it is classified by an automatic classification system. This classification is validated by an editor that may change it if necessary. The author will be notified of the classification given to the document. Afterwards the document may be sent to a peer review process or not, accordingly to the genre of document in hand.

The peer review process is blind, but conversations between different intervenients are allowed; at any moment these conversations may become public, if this is the editorial board desire. Reviewers are asked to fill in a classification form in order to guarantee some coherence in the review system among different documents and different reviewers. Reviewers are classified by editors and, if necessary, may be invited to leave the reviewers body (e.g. systematic delays for reviewing an article should not be allowed). Compensation and promotion mechanisms for reviewers will have to be considered.

**Editing and Publishing**
After acceptance, documents will be kept in an editing period where reviews can still be made by the authors and the editors. In this period, layout and file formats issues are also addressed by the publishing team. It is here that, for the last time before publishing, the document’s hypertext links and other references are checked. File formats will be chosen for storing the document on the database. If the reader requires another format for downloading, online or offline reading, a file format converter system will be used - this information is going to be stored in her/his profile.

**Reading and Commenting**
Reading a document is much more than just reading in the common sense. First of all, a reader can choose the format (in a list) and the language in which to read or download a document. For downloading, he/she can choose different formats whether it is for offline reading or for paper reading. These formats and the default language are predefined in the user’s profile, but they can be changed (permanently or not) at any moment. The reader is also allowed to make a document list for online viewing, downloading or printing. Each of these documents can have different file formats. Each document or section of document is presented with hyperlinks to the commentaries made to it, which the reader is allowed to see. The reader is also allowed to make generic searches or customized searches that are made accordingly to his/her profile. The result of these searches can be shared with other users.

Generally, comments will not be anonymous. The pre-condition for commenting is to be
identified and identifiable. Anonymous comments are allowed as long as they go through a review process, not for censoring content, but for censoring form. A commenter may comment the entire journal (in which the comment is transformed in a letter to the editor in chief), a document, a section of a document or a comment. In commenting a document or a section of a document, the commenter may make reference to other documents in the database. This comment will be attached to the other documents also.

Collaboration

The collaboration module is not going to be described because it will not be implemented for now. We will try this system and study its acceptability and the information behavior of intervenients. We will also study hidden collaborations triggered by the system. Only in a later phase, the collaboration module may be implemented.

Research

Additionally to these modules we are designing the requirements of the research module. We think that this is a very interesting module because it will serve only for research purposes: it will have a kind of a monitoring system of the users behavior in order to study and draw some scientific acceptable conclusions on this matter.

4.3. Users’ profiles

All these modules will work being aware of the user’s profile. There will be several kinds of users, or better, several kinds of roles a user can perform: author, reader, commenter, reviewer, editor, editor in chief, publisher and editor assistant. Each user will be able to perform several roles whose information will be stored in her/his profile.

Fig. 4 - Class inheritance hierarchy for roles of Informattica Online users
Roles have common characteristics and can be represented using a class inheritance hierarchy as shown in fig. 4.

The users’ profile relates not only to the contents he/she is interested in, but also to the form in which those contents are accessed or retrieved, and to other user characteristics such as the kind of user he/she is. It will have a common part with general information and specific parts according to the role the user is playing. In the user’s profile there shall be information such as the user’s interests, ranking criteria of documents, formats used for online, offline and paper reading, formats used for media downloading, information about the default role (the way the user wants to enter the system), and so on.

5. Future Work
We are using Rational Rose to design the system in UML (Unified Modeling Language). For now we have designed part of the requirements of the system with Use Cases. Neither the Research nor the collaboration modules have yet been designed. The latter will be left for designing after the system is implemented and tested: first we want to draw some scientific acceptable conclusions on the system’s usage.

After completing this phase, we will go through the system’s analysis and design processes using Rational Rose. Concerns with metadata for new document genres (or for new processes over those documents), the handling of multiple versions and of dynamic documents, relationships between formal and informal documents and related processes, individual work sharing, citation issues (for formal and informal documents), profile construction and personalization are all issues that will be taken into account. On the other hand, there are some subsystems that will be bought and others that will be object of collaboration with other related work teams (like the automatic classification system, the multi-language translator system, the search engine, etc.).

The Portuguese informatics scientific community will be subject of a study regarding its information behavior related to the use of scientific journals. The way people read journals, acquire information, interpret it and use it on their work and daily activities will be carefully studied. After the system is implemented, the same study will be made regarding the information behavior using it.

6. Conclusion
With Informattica Online we hope to provide an environment for scholars to communicate effectively, even if they are in distinct fields. By maintaining a large base of documents we cover several fields of knowledge; by providing personalized access and services we respond to each user personal needs. Common access to same documents of people coming from different areas is expected to generate interdisciplinary communication. Part of the undiscovered public knowledge may also be found.

We try to take many of the advantages that technology provides. However, we also try not to forget what a journal in fact is and might be. Formal communication documents are the basis of Informattica Online; a high quality peer review process will be implemented. Nevertheless, we complement these with new genres of documents, metadata and processes that take real advantage of the electronic medium.

References


Notes
1. Assistant Lecturer and PhD student at the Department of Informatics at Universidade do Minho, Portugal.
2. Librarian. Head of Information Services Division at Universidade do Minho, Portugal.
3. Full Professor at the Department of Informatics at Universidade do Minho, Portugal.