The Design and Development of Internet and Intranet Based Collaboratories

Anitesh Barua
Ramnath Chellappa
Andrew B. Whinston
barua,ram,abw@cism.bus.utexas.edu

Center for Information Systems Management
Department of Management Science and Information Systems
Graduate School of Business
The University of Texas at Austin
Austin, Texas 78712

Tel: 512 471-7895, 512-471-7962, 512-471-8879
Fax: 512 471-0587

We are grateful to the Editor and anonymous reviewers for many helpful comments and suggestions.

The Design and Development of Internet and Intranet Based Collaboratories

Abstract

In today's interdependent world, applications supporting collaborative interactions are critical for individuals, groups and organizations alike. Most groupware applications are based on proprietary standards, necessitating vendor-specific tools and making it difficult for *ad hoc* collaboration. This calls for a shift of technology platforms for collaborative systems ("Collaboratories") from proprietary to open environments based on the Internet and Intranets. We enumerate and analyze Collaboratory requirements for three geographically dispersed electronic communities: individuals with overlapping interests but without formal ties, special interest groups, and organizations with interdependent workgroups. The requirements analysis is based on the premise of maximizing the net value derived by Collaboratory users, and identifies sets of infrastructure and user community specific characteristics. We show to support these requirements, there needs to be an integration of Internet applications such as the World Wide Web and USENET and non-Internet technologies such as database systems and scripting languages. We also discuss some design philosophies for preserving the open but secure (where appropriate) nature of collaborative systems. We follow up our requirements analysis with the implementation of three Collaboratories: one based on the Internet with a flat-file structure, one involving Web and database integration in an Intranet setting, and one based on both the Internet and an Intranet. We also report the results of a survey of users of the first Collaboratory, which support our theoretical premise of a technological platform shift and the need for integration of multiple technologies for effective collaboration.

Keywords: Electronic communities, collaboration, the Internet, Intranets, newsgroups and the World Wide Web.

1 The Increasing Importance of Collaborative Interactions

In today's "interdependent global village" [19], the need to interact, collaborate, communicate, and learn across the globe is becoming ever more critical for various communities such as informal groups of individuals with overlapping interests, groups and societies with special interests and expertise, and both commercial and non-commercial organizations, which have been described by various scholars as "dispersed occupational communities" (e.g., [29, 24]) and collections of "communities of knowing" [7]. For individuals and special interest groups (SIGs), bulletin boards and USENET newsgroups on the Internet have long been a platform of interaction regarding a variety of issues and topics. Individual students, academics and professionals have always found it useful to retrieve information relevant to their professional interests, to discuss problems and ideas with peers and experts around the world, and to make announcements to groups of people with common interests. Organizations too are discovering that dispersed team interactions both within and outside the organization are a key aspect of their daily operations in terms of coordinating workflows [19, 14, 22] involving multiple agents leading to successful outcomes. The widespread adoption of electronic commerce (EC) is also necessitating high levels of communication, coordination and collaboration. As a result, the importance of technologies and applications to support collaborative interactions has increased dramatically in recent times.

Adapting a definition by Barua, Chellappa and Whinston [2], we consider a collaborative system or "Collaboratory" as an open electronic platform for individuals, SIG or organizations to efficiently exchange, disseminate and create issues, ideas and knowledge, and to accomplish *ad hoc* tasks towards the fulfillment of shared objectives. We first make the case that to cater to the collaboration needs that arise in an *ad hoc* manner in different communities, collaborative systems must make a transition from proprietary technologies to Internet- and Intranet- based systems.

Being based on non-proprietary standards such as the Transmission Control Protocol/Internet Protocol (TCP/IP) and associated applications accessible to users on a global basis, the Internet has witnessed a rapid growth of interaction tools such as USENET, Multi-User Dungeons (MUD), Internet Relay Chat (IRC), and the World Wide Web. The Web provides the ability to display multimedia documents with hypertext links, making a quantum leap in information access and presentation capabilities over Wide Area Networks. However, we argue that a collaborative system requires a combination of various features and functional capabilities, and that no single Internet application in isolation supports the complete set. Thus, there is a void in terms of an integration of various Internet and non-Internet technologies, which we consider a critical prerequisite for Collaboratory development. Similar views are reflected by Applegate and Gogan [1] in a more general context: "The connectivity, flexibility and power of the tools available on the Internet provide an excellent platform for delivering interactive real-time information both inside and outside the firm. But taken alone, they are insufficient. The Internet's power must be integrated with the internal databases, networks and transaction processing systems within each firm that unites to deliver value chain activities."

While Internet based Collaboratories may serve the interaction needs of individuals with common interests, organizational Collaboratories must reflect the underlying business protocols and their implications (e.g., security, workflow, etc.). Thus, the latter must be constructed within an Intranet environment, which provides the necessary infrastructure for these additional capabilities. Contrary to the popular trade press notion of Intranets as internal corporate web servers, Chellappa, Barua and Whinston [8] envision an Intranet as "a secure corporate network with rich functional features of Local Area Networks (LANs) interconnected by the Internet and/or its technologies and applications." In other words, an Intranet is a suitable platform for building applications that support security, reliability and functional demands placed by business rules and processes.

We enumerate and justify certain characteristics or features of a Collaboratory which are geared toward maximizing the value (benefit obtained from a Collaboratory minus the opportunity cost of use) derived by users belonging to various communities. After a high level requirements analysis, we focus on design philosophies and the integration of various technologies (e.g., the Web, databases and scripting languages) we consider essential in building Internet and Intranet Collaboratories. We first describe our implementation of a Web based MIS Collaboratory (in 1994) using a flat-file structure. We also present the results of a survey of the Collaboratory users conducted between September and December 1995. The empirical findings support our position regarding the Collaboratory requirements, and show that Collaboratories of this type have a high potential for SIGs like the MIS community. Further, our conjecture regarding the need for Web and database integration to enhance collaborative interactions was also strongly supported by the survey. This motivated us to build a database-enriched Web platform, and to develop two additional Collaboratories on this foundation. One was designed in an Intranet setting for a SIG within a University, while another was developed as an on-line journal supporting both private and public interactions to expand the frontiers of electronic commerce.

The contribution of the paper is a rich conceptual basis for the motivation and requirements for open collaborative systems, multiple implementations (two of which required an integration of Web and database technologies), and empirical findings which can guide future developments in this area. Apart from demonstrating the role of open technologies in the development of collaborative environments, the Collaboratories we have designed and implemented provide a starting point for research on systems supporting EC activities. Intra and inter-organizational collaboration is an integral component of EC, and our study offers guidance on how such collaborative activities can be supported within the context of business rules and protocols.

2 Making the Case for a Platform Shift

While the need to form geographically dispersed *ad hoc* teams in organizations and to link them electronically is rising sharply [18, 19], organizations have primarily relied on proprietary technologies and platforms (hardware, operating systems and communication standards) to provide the required capabilities [2]. For example, electronic mail based on proprietary technologies have been extensively used in organizations for nearly three decades [17, 24]. Even recent groupware applications such as Lotus Notes are based on proprietary standards; while Notes can now work over the Internet (due to expensive "Internotes" gateways), the end-users must all have Lotus Notes to be able to communicate with each other [3]. Such systems are not open in the sense that it is not possible for users with different applications to interact on a global basis. While these systems may cater well to the needs of individual organizations or workgroups, they appear less suitable for supporting potentially global interactions among people engaged in the collective pursuit of advancing a specific objective.

Our economic argument against proprietary platforms and applications is that they require *asset specific investments* [32], which cannot be readily re-deployed in other settings, force an organization to commit itself to a specific vendor's products and upgrades, and make it difficult for new trading partners to become a part of networked or virtual organizations (unless they use the same vendor's products). Apart from the connectivity and interoperability problem, another technical shortcoming of using vendor-specific applications is that users are completely dependent on one vendor in providing the entire range of functions and capabilities. It is definitely advantageous to be able to incorporate various technologies developed by a variety of vendors or individuals, provided they can work seamlessly. For example, a Web-based SIG Collaboratory we have implemented (described later) supports voting sessions. Currently, voter registration and password protection are the only integrity/security measures for the voting workflow. However, we can even integrate vendor specific authentication systems within the Collaboratory, provided the former are based on Internet protocols.

By virtue of being available on a world-wide basis with open client applications like Mosaic and Netscape, the Web appears to be ideally suited as the Internet technology on which a Collaboratory can be developed. Depending on the document type, Web browsers seamlessly invoke other applications such as word processing and spreadsheets. Thus, a user can make available on the Collaboratory a spreadsheet file where s/he has conducted some analysis to support his/her position on an issue. Other users can use browsers such as Netscape to invoke the spreadsheet application, modify the analysis, and obtain a different set of results. Thus a Web based Collaboratory will enable users to reference and access virtually any type of document created by commonly available applications. In addition to these capabilities, an Intranet setting will help capture the business protocols of an organization or SIG, and therefore appears to be a promising platform for organizational and SIG Collaboratories.

3 Requirements Analysis for Collaboratories

3.1 Interactions in three electronic communities

We consider three communities as potential users of the Collaboratories we envision: (i) informal groups consisting of individuals with partly overlapping interests (e.g., people with an interest in foreign cultures) but without formal ties, (ii) societies or formal groups with special interests (e.g., MIS), and (iii) commercial and non-commercial organizations. The distinction between these communities lies in the extent to which members of a community interact for satisfying personal interests versus achieving professional goals, and the structuredness of their interactions. In order to enumerate Collaboratory requirements, it will be useful to characterize the three communities in terms of the nature of their interactions and traditional applications that have supported these objectives.

Interaction amongst individuals is largely informal, with little or no structure governing the process except for those imposed by the interaction tools themselves. Such tools include Internet newsgroups, chat rooms, and MUDs. It should also be noted that much of the interaction involves information gathering and dissemination, and taking positions on issues (generally without formal argumentation and discourse).

By contrast, a SIG requires a more structured environment for meaningful interactions restricted to a particular domain of interest. Naturally, there is a greater degree of formalism to these interactions, and a collective pursuit of a shared objective (e.g., advancing the frontiers of knowledge in a discipline). For example, consider ISWorldNet (*http://www.isworld.org/isworld.html*), a SIG consisting of IS academics and researchers from around the world. It involves standardization of document formats, appointments (such as editors for various sections) and formal channels of communication among office-bearers. Later in the paper, we will argue that the technology requirements for such a group are quite different from those for individual users. Traditionally, mailing lists (LISTSERV) have served as the medium of collaboration for SIGs.

Interactions between individuals, workgroups or teams within or among organizations are characterized by firm-specific goals (e.g., meeting deadlines for project completion, new product introduction, etc.), and strict adherence to business rules and protocols (e.g., decision authority, confidentiality, etc.). In recent times, proprietary groupware technologies such as Lotus Notes have supported these capabilities.

3.2 User value maximizing characteristics of Collaboratories

Qualitatively speaking, the net value derived by a Collaboratory user has two aspects: the benefits resulting from interactions, and the cost of time and effort spent on collaborative systems. Benefits are derived by accomplishing objectives (e.g., finding an inexpensive air ticket for an individual, getting key references on a particular topic in a SIG context); the nature of costs incurred depends on the user – for an individual, it is just the cost of time (and associated boredom and frustration) spent on browsing newsgroups and posting or soliciting information. For an organization, this cost can involve the integration of diverse applications and technologies to enable collaboration with a trading partner using an incompatible system.

The above characterization helps us assess requirements which may be more salient for one community than for another. We propose certain characteristics as drivers of the benefit and cost of using a Collaboratory, and classify them as either *common infrastructure characteristics* (if they apply equally well to all three communities) or as community specific (if one or more communities need such characteristics more than others).

3.3 Common infrastructure characteristics

Based on primary goals of Collaboratory usage (as highlighted in the definition of a Collaboratory), we propose the levels of *information access*, *interaction richness*, and *information scalability* as key common infrastructure characteristics.

3.3.1 Level of information access

One of the major objectives of using a Collaboratory is information gathering, whereby a user can seamlessly access *relevant* information dispersed throughout the world of connected networks or an organizational Intranet (where appropriate). The overall architecture is envisioned as a set of central points distributed throughout the globe that serve as gateways to a specific field of interest and related information. The search among these distributed Collaboratories would be made more efficient and effective through the appropriate organization of information resources, the use of context-specific search mechanisms, and the Web's dynamic linking capability [2]. Information resource organization would involve categorization of issues, topics and ideas [31, 9], as well as virtual views of "meta-information" resources [23] potentially located across the entire Internet and multiple organizational Intranets.

3.3.2 Level of interaction richness

A Collaboratory must support value enhancing interactions with other members of an organization or a society to debate issues and problems of importance and interest, as well as to gain new insights, information and knowledge. Interaction richness can be defined as the extent to which the barriers of space, time, and media/document formats can be overcome in interacting with others over a wide geographical scope. This would include the ability to talk, see, write and draw in both synchronous and asynchronous manner, access to relevant reference information, archiving of interactions for future review. This notion is related to media richness (see [33] for operationalization and measurement of this construct) in a Computer Supported Collaborative Work environment, which has been defined as the ability to convey expressions, pitch, tone, etc. [12].

Whether the ability to mimic face-to-face meetings through a Collaboratory actually adds to or detracts from the value of the Collaboratory depends on the context of use (e.g., see [5] for an analysis of both positive and negative impacts of anonymity in group support systems), our goal is to provide an array of features, which can be selectively chosen or left out depending upon contextual factors of an interaction environment. The design features that provide a high level of interaction richness can be grouped under the following categories: *dynamic linking capability*, *synchronous* and *asynchronous communication*, and *multimedia support.*

The ability to link remote documents and resources "on the fly" is critical to interactions based on sound arguments, rational and empirical/anecdotal evidence. For example, in a discussion of the information superhighway on an electronic forum, a user may take a position that the Internet is only a small subset of the superhighway. In an ideal electronic forum, the user should be able to create links to supporting documents. The ability to interact in both asynchronous and synchronous modes is crucial to group interactions and should clearly be an integral part of a Collaboratory. Groupware applications have achieved this capability within LAN environments. Internet applications such as Newsgroups and Internet Relay Chat also feature these abilities. The Web by itself does not support such interactions, since it was primarily designed for the dissemination of information stored in the form of documents. However, in contrast to the other Internet tools, the Web does support a wide variety of multimedia documents, another key element in creating a realistic collaborative environment.

3.3.3 Information scalability

From a pure technology standpoint, scalability refers to the ability of a computing system to meet performance requirements as the load or the size of the system increases. However, we introduce the notion of *information scalability* as the ability of a Collaboratory to avoid information overload on a user as more information gets added, and as the volume of interaction increases. Instead of transferring the increased information processing burden to the user, the collaborative system should manage, filter and present only relevant information to user. This is not just a re-organization of information, but involves the dynamic customization of content based on user needs. While information organization (as depicted in section 3.1.1) does not cater to specific user needs, by capturing user profiles and preferences, an effective filter can be created. For example, instead of presenting users with the entire content of a new discussion issue, pointers to the new topic can be sent as an automatic notification. Hence, an user who has registered with, say, a marketing section of an organizational Collaboratory can choose to be informed of any new postings only in that particular area. By contrast, in a USENET environment, the user manages what s/he views/reads, and needs to browse through a vast amount of information. While a user can partly customize his/her view by subscribing only to select newsgroups, this level of customization may be insufficient, given the volume of potentially irrelevant information appearing in the newsgroups.

3.4 Community specific Collaboratory characteristics

As shown in Table 1, Collaboratory characteristics which are specific to an electronic community grow in importance (as indicated by the number of $\surd$ marks) as we move from informal groups to collaboration in organizational settings. This theme is developed in the balance of the section.

******INSERT TABLE 1 ABOUT HERE*****************

3.4.1 Intellectual organization of archived information and interaction history

Internet newsgroups have no organization of information in terms of a shared paradigm or understanding of a related set of issues. This is not a limitation in a newsgroup setting, where there are no common objectives to be achieved collaboratively with immediate cost and benefit implications. However, a SIG (e.g., the MIS community) will need more focus and structure in storing information and interactions on a Collaboratory, because there is at least an informal shared goal of advancing the frontiers of a specific field. For organizations, the need to structure and organize archived information and interactions based on key business processes and their information requirements is clearly a critical aspect of collaborative work.

As an illustration of information organization based on a shared understanding within a special interest group, consider a Collaboratory for MIS academics, professionals and students. Such a Collaboratory should be a direct reflection of a shared conceptualization of the field of MIS. Barua, Chellappa and Whinston [2] argue that MIS creates new knowledge and value for both academics and professionals by combining technological and managerial viewpoints. The managerial viewpoints come through multiple reference disciplines such as sociology, economics, organizational behavior, marketing, etc. In other words, the technological and managerial viewpoints are complementary, and new knowledge is created at the intersection of Information Technologies (IT) and the reference disciplines [2]. This can provide a basis for organizing information in a Collaboratory.

Note that the organization and structuring needs discussed here is different from the information access issues outlined under common infrastructure characteristics. While the latter deal with improving the overall efficiency of information access through features such as indexing, unified virtual views of widely dispersed information resources, the former deals with the intellectual basis within which interactions take place.

3.4.2 Validity and reliability of information, knowledge and user inputs

One problem with applications like Internet newsgroups is the trustworthiness of the information posted. There is little verifiability of either the messages or the sender. While such applications may serve the needs of individuals with similar interests but without a common objective and a shared commitment to an organization or group, a Collaboratory for the latter must support the process of establishing content validity (the extent to which information or messages are accurate and qualified) and the reliability of the users.

One of the earliest approaches for analyzing complex problems as a structured conversation between its stake-holders is the Issue Based Information Systems (IBIS) method [20, 27]. Key elements of IBIS include issues, positions, arguments with various links between the elements. It helps the stake-holders understand others' positions, but does not provide mechanisms to manage argument based interactions. Ramesh and Whinston [26] extend the IBIS framework by developing the notion of Argumentative Reasoning Facilitation Systems (ARFS), based on formalisms not only for representation of argumentative knowledge, but also for conducting, regulating and coordinating the interaction process. A Collaboratory for a SIG or an organization can benefit significantly from a Web based implementation of the ARFS system, simply because views, positions, decisions, etc. will be subject to argumentation and qualification before they are considered valid or acceptable. This feature helps create what Boland and Tenkasi [7] describe as electronic forums for *interpretive reading* and *theory-building*. Of course, we envision these representation and validation capabilities as an integral part of an overall Collaboratory, rather than as isolated systems.

While the validity of user inputs can be established through dialectic schemes such as the one described above, a Collaboratory can also take advantage of users' reliability levels and reputation in providing incentives for users to contribute quality inputs. Ching, Holsapple and Whinston [10] consider how an organization can use reputation of its members as a capital which appreciates or depreciates based on the members' actions to improve performance. A Collaboratory can maintain indicators of users' reliability and reputation in a variety of ways. First, annotations and positions in favor of and against a user's inputs can be maintained automatically. User inputs and interactions can be archived as part of organizational memory [30], which can be retrieved by other users in different contexts. The frequency and subsequent use of this memory can be deployed in measuring reliability and reputation indices of a contributing individual.

3.4.3 Process and workflow support

An individual newsgroup reader is not part of any business process or formalism that would typically characterize the activities of SIG (e.g., nominations, voting, policy making) and organizations (e.g., operational and strategic business decisions). Georgakopoulos, Hornick and Sheth [15] define workflow as "a collection of tasks organized to accomplish some business process." The business literature on workflow classifies processes into three types [15]: *ad hoc*, administrative and production. The type of workflow support we envision for a Collaboratory do not involve routine tasks such as order processing; rather, we are interested in *ad hoc* processes like voting, new product decisions, and electronic review of articles, which involve selection, negotiation and collaboration. As we have articulated earlier, process support comes with its security/integrity needs as well as the integration with a variety of non-Internet technologies, and necessitates Intranet based implementation.

3.5 Complementarity: A theoretical rationale for coordinated choice of characteristics

Based on emerging theories of complementarity in economics [21] and its application in MIS in the form of business value complementarity [6], we suggest that all the above capabilities must be provided in tandem through a single integrated application. Complementarity implies that the value of such a convergent system to a user will be more than the sum of the values of isolated systems each of which provides one of the characteristics discussed earlier. Theoretically, it is possible to enjoy many of the above characteristics on the Internet in an isolated manner. For example, Web sites already have a large volume of relevant information on virtually every topic. The Web also supports multimedia documents. USENET bulletin boards allow asynchronous interactions on a global basis. Internet Relay Chat provides synchronous interaction capabilities. Wide Area Information Servers (WAIS), Archie, and their derivative search applications allow users to search for documents and their locations. Collaborative interactions, however, require all these and other capabilities. From a user's perspective, the cost of the time and effort involved will be too high if separate applications are needed for each aspect of an interaction.

Incorporating these capabilities within one system can significantly reduce the users' cost of spending time on a Collaboratory. A traditional Web, which has multimedia capabilities, cannot by itself support asynchronous interactions like newsgroups. On the other hand, newsgroups lack the multimedia capabilities of Web-based applications. Clearly, the benefit of having both features together is very high from the users' standpoint. Similarly, the ability to dynamically link to an information resource to support one's stand on an issue is a key feature of a constructive interaction. While the Web allows hypertext links, it does not come with a platform for asynchronous or synchronous communication. Once again, the benefit of having the platform and the linking capability together is higher than the sum of the values derived from the two separate systems, one with the linking capability and the other with the forum. Or consider the interplay between multimedia support and dynamic linking capability. Being able to dynamically link to, say, FTP sites, but not retrieve diagrams or pictures associated with a document will be much more tedious (and hence costly) than using a system which allows both dynamic links and multimedia documents. In essence, we are arguing that the "whole is greater than the sum of its parts". Undertones of complementarity are also echoed in Neuman's [23] pioneering research on Internet information resource organization: "There are four areas where work is needed to help users obtain the information they need: retrieval, indexing, search and organization. A number of recent systems have addressed the first three areas, yet the fourth has been greatly ignored. Users require all four functions if they are to obtain the information they need. Without work on organization, the other functions become less useful as a system grows."

4 Technology Issues and Design Philosophy

Being just a presenter of information, the Web by itself cannot be a complete platform for a Collaboratory. We find that a wide range of technologies must be integrated to provide such support: open systems (such as the TCP/IP standard) for easy access without specific system requirements, relational database management systems for providing only relevant information for users and for efficient storage of information, the Web for multi-media and hypertext link support, and cryptography for security. To elaborate a little, database management systems have proven themselves to be the most efficient way of handling data, while the Web has established itself as an ideal front-end for the transfer and display of multimedia documents. It may be therefore be efficient to let a database store Collaboratory data in terms of records and fields, while the same information will be selectively presented by Web front-ends as dynamically created HTML documents. The database would also enable us to customize a Collaboratory to an Intranet setting, where business protocols involving privileges, activity sequences, etc. can be easily enforced, and where different users can specify their profiles and preferences to be presented with information that appears to match their requirements. The integration will be achieved by embedding applications at the server level so that the Web client would do no more than request documents. Common Gateway Interface (CGI) scripting is one way to implement such a scheme.

Our key design philosophy was to maintain the open nature of the Web. We felt that no changes should be required at the user (client) side (i.e., any HTML 3.0 browser should be adequate), and that no proprietary protocols or features should be added. This will enable any authorized user with just a Web browser to access a Collaboratory. Another reason why we did not want to "empower" the client side involved the general security of the Collaboratory. Having minimal client functionality (e.g., the ability to only send form inputs and to display documents) prevents hackers or unauthorized users (in the case of Intranet based Collaboratories) from getting access to the server with an ulterior motive.

5 Implementation Details

We proceeded with a phased implementation approach. Our initial objective was to exploit the rich multimedia and resource linking capabilities of the Web in a groupware setting, and to provide an integrated technology platform (Web + database + scripting) for a Collaboratory.

5.1 A flat-file Collaboratory

Our attention was focused on how users could dynamically create HTML documents on our server without any manual intervention. Since an HTML document is just a regular ASCII (text) file, it was natural to begin the implementation of the Collaboratory as a flat file system. We developed CGI scripts using an interpreted language called Perl to handle file and directory creation. CGI scripts allow HTTP requests to execute processes on a remote server through the use of a normal browser and form inputs. We used these scripts to manage both information dissemination and interaction on the Collaboratory. Consider an MIS professional adding a comment to a discussion or making an announcement regarding an upcoming conference. In either case, s/he fills out an appropriate form and clicks on a send button, which activates our CGI script. The latter deals with this document posting request and evaluates it to determine the appropriate directory to place the document. Once the document is placed in the appropriate repository, it sends a message to the document displayer unit to update the index document on conference announcements. Thus, both the information positioning and dynamic indexing are handled by this unit. This is shown in Figure 1. Based on this approach, in 1994, we made available on-line what to the best of our knowledge is the first interactive Web-based Collaboratory (*http://cism.bus.utexas.edu/issues/iindex.html*). Currently, this Collaboratory is used by MIS professionals and academics from all over the world, and on an average over 200 different issues are discussed and debated at any point in time.

**************INSERT FIGURE 1 ABOUT HERE**********

5.2 Survey of users of the flat-file Collaboratory

After the Collaboratory was in use for about one year, we conducted a user survey between September and December 1995. The objective of the survey was to assess (i) the perception of the users regarding the usefulness of Web based Collaboratories relative to newsgroups, and (ii) the perceived importance of various features (existing and planned) of our Collaboratory. This survey was administered through the Web. Since the objective of the survey was not to collect general information on the Internet or Web users, there was no need to conduct a non-web survey.

While there are no well-established guidelines for correction factors in Web-based surveys, there is a major threat of unequal visibility to various members of the target group [16, 25]. We addressed this concern in the following manner. Starting in 1994 (when the Collaboratory was developed) we have announced its existence and purpose multiple times on all relevant newsgroups, mailing lists, and registered the site with Internet search engines as an MIS Collaboratory. This ensured that members of the MIS community on the Internet had an equal chance to use the Collaboratory.

It was important to check that the survey was completed by respondents who had actually used the Forum and not by any Web or Internet user. We wrote a special script to capture all accesses to the survey with built-in features to eliminate consecutive accesses from the same machine (assuming reloading of the document) and accesses by search engines (such an access implies the survey was indexed by some archival facility, and not really seen or used by any individual). Thus, we were able to capture the number of users who actually looked at our survey. Further, the script checked for respondents who might have answered the survey questions without using or browsing through our system. These responses were ignored.

Since the MIS Collaboratory or the survey required no special browser capabilities, we believe no bias was introduced in the form of special technology requirements. Also, every page in the Collaboratory had a link to the survey, thereby making the survey visible to all users.

A total of 1237 Collaboratory users (referred to as users from now on) looked at the survey, while 281 users completed the questionnaire, resulting in a response rate of 22.72%. Although there are no pre-established factors for Web-based responses, this falls in the same range as mail surveys [28]. The profile of the 281 respondents shows a mix of academics, students and professionals (see Figure 2). A majority of the respondents (68%) have been using the Internet for 2 years or less (see Figure 3), possibly indicating that they started using the Internet after the advent of the Web in 1993. Interestingly, at the time of the survey, a sizable fraction (74%) of the respondents were using the Internet often or very often for work-related activities (Figure 4). Overall, the profile of the respondents was representative of our target population of serious Internet users. We are clearly less interested (as far as SIG and organizational collaboration is concerned) in casual Web surfers who use the technology primarily for entertainment.

*******INSERT FIGURES 2, 3, 4 ABOUT HERE***

The respondents were asked to rate the Collaboratory for its usefulness in supporting MIS related interactions. A five-point scale ranging from "not at all" to "very much" was used. 46.62% respondents felt that the Collaboratory was "quite a bit" or "very much" an effective medium for MIS related interactions, while 40.57% felt it the system was "somewhat" of an effective medium (see Figure 5). Since newsgroups were the primary means of collaboration on the Internet during the time period covered by the survey, we asked the respondents to rate the Collaboratory's usefulness relative to newsgroups as a platform for constructive interactions in the field of MIS. 64.06% respondents rated our system as "quite a bit" or "very much" better than newsgroups (see Figure 6).

********INSERT FIGURES 5, 6 ABOUT HERE *************

We also asked the respondents to rate the perceived importance of a variety of existing and planned features of the Collaboratory. Two of the key features, the ability to post issues and comments automatically and to dynamically link other documents, were judged to be very important by the respondents (see Figure 7). Interestingly, the correlation between these two items is .40 ($p < .0001$), indicating that respondents who considered automatic posting ability as important were also likely to find dynamic linking an important feature. Indeed, our theoretical premise of complementarity suggested that the joint impact of these features would be more profound than that of having them in isolation. It lends credibility to our design philosophy that all attributes must be provided in an integrated fashion.

The respondents were somewhat less excited about the ability to use images and voice annotation on the Collaboratory. For example, 43.42% (55.87%) felt that it was "not very much" to "somewhat important" to embed images (voice) in the interactions (Figure 7). We can speculate on different reasons for this lack-luster ratings. First, the type of interactions we witnessed on the Collaboratory primarily involved ideas, positions and arguments, which did not require substantial multimedia support. By contrast, interactions on a Collaboratory for a field such as Astronomy or Archaeology is likely be visually intensive. Also, limited modem and line speeds may be an important factor influencing the respondents' perceptions of these bandwidth-intensive features. The correlation between the importance of image and voice was .49 ($p < .0001$), indicating that users who considered one of the features as important also considered the other as important. It appears that some of the Collaboratory users were multimedia enthusiasts, and considered features such as images and voice annotation as important.

********INSERT FIGURE 7 ABOUT HERE *************

For additions to the flat-file system features, 75.44% considered enhanced search capabilities similar to those provided by database systems as "quite a bit" to "very much" important (see Figure 8). Also, 62.64% felt that there was a pressing need for managing discussions (along the lines discussed in sections 3.3 and 3.4). It is interesting to note that efficient management of interactions is not feasible on a Web based system without the integration with database technologies. The correlation between the importance of database search and interaction management is .39 ($p < .0001$), again indicating the importance of having both features together.

37.72% suggested that voting capabilities on a Collaboratory is "quite a bit" to "very much" important. This is not surprising given that the audience forms a SIG in the MIS domain. For a larger SIG such as ISWorldNet, this capability would appear to be even more important. Only 10.32% saw the need to integrate synchronous conferencing tools like MBONE, Cu-SeeMe with the Forum. This feature would certainly be much more important in organizational settings.

********INSERT FIGURE 8 ABOUT HERE *************

The above results were largely supportive of our characterization of the technology requirements of a SIG. First, it confirmed our premise that semi-formal or formal collaborations would find a Web-based forum more useful than newsgroups. The need to organize discussions (which could not be done efficiently with a flat-file system) and provide enhanced search capabilities (in the flat-file implementation, we had a simple WAIS-based search feature) also became apparent. The users were also favorable towards establishing workflow processes such as voting mechanisms. As discussed earlier, an integration of the Web and database systems would help fill these gaps.

5.3 Integrating Web and database technologies

Having achieved the initial goal of Web interactivity, we found that automated organization and management of data created by the users across the globe was inefficient with a flat-file system (this was also confirmed by the above survey). Additionally, it was not possible to have flexibility of presenting information based on user profiles and requirements. To address these limitations, we integrated the Web and database systems (see Appendix for details on Web-database connectivity) by writing CGI scripts to handle data querying on the database side (Oracle V.7.1.6), and data formatting for the client side. The conceptual design is shown in Figure 9. A user logging on the electronic forum for the first time sees a login screen where s/he is required enter his/her login name and password. This input is transmitted by the local browser to the Web server. The Web server forwards the user's input to the SQL query generator we have developed. The latter parses the request to create a SQL query, which is then passed on to the database engine to request data from the relevant database tables. Since this was the first time the user logged on, the SQL query requests a list of discussion topics. The data returned from this SQL query is passed back to an HTML document formatter, another Perl script that we have created. This formatter dynamically creates an HTML document out of this data, and passes it back to the user. To the user, it appears as if s/he is accessing a regular HTML file. But note that the entire forum is now based on the principle of true dynamic document creation. In other words, no HTML documents are stored, and are created "on the fly" in a transparent manner. We argue that this database-enriched Web presents a new platform for the development of Intranet based Collaboratories, which require a broader range of capabilities that are best handled by database technologies.

**************INSERT FIGURE 9 ABOUT HERE**********

5.3.1 Customizing the database-enriched Collaboratory to an Intranet environment

The New Initiatives in Digital Sciences (NIDS) Committee was set up at the University of Texas at Austin to build awareness of the "Digital Revolution" related issues, challenges and opportunities facing the University. The committee consists of distinguished senior faculty members from various departments. Their use of a Collaboratory would include discussing and debating issues, setting committee agendas, archiving minutes of meetings, and engaging in voting sessions. These interactions were to be kept confidential from the public domain, and would hence have to be implemented in an Intranet setting. However, since one of the NIDS objectives was to attract sponsors and donors for a new institute, some of the information being generated on the Collaboratory would be made available to a select external group, while some other types of information would be completely in the public domain after the committee had engaged in private Intranet-based interactions. These requirements were met by customizing the database-enriched Collaboratory to include appropriate security measures. Additionally, the database allows many enhanced features, including the dynamic customization of messages, automatic electronic mail notification regarding new postings of interest, and selective deletion privileges for NIDS members.

The database was organized in an object-oriented fashion with a "table-to-inherited type" (or a type-tree) structure [13]. Flexibility of extension and modification is a major advantage of such a mapping. For example, if the Collaboratory were to be integrated with another pre-existing database, it would be easier to extend the table-types to fit the latter structure. Further, handling data-types such as audio or video files are greatly simplified by this approach. In a traditional database modeling approach, a purely relational schema is pursued, and binary "objects" such as images files are not handled efficiently. By integrating object-oriented views of the database, the scripting part of the system (which handles the formatting of HTML documents), becomes more efficient. In our system, instead of storing image/voice files as a "long raw" field in the relational database, we simply store a reference to the object. Therefore, objects associated with their parents may be easily inherited as a reference rather than querying the database again.

The present implementation consists of four tables: user, issue, comments and votes. Every user of the forum is uniquely identified by his/her login id (primary key). As soon as the user logs in, all attributes of the user such as email, issues started by the user, and the voting sessions initiated by the user are queried from the database. The resulting information is inherited by every action invoked by the user, thus giving the ability to customize each user based on his/her profile (attributes). Figure 10 shows a partial and simplified schema of how a class is constructed from different tables. Each issue is uniquely identified by the issue number (primary key); subsequently the comments are referred to by a combination of the issue number and comment number (foreign key). The text part of the issues and comments are stored as references to files. Each user is allowed to initiate a voting session, and the object-based model allows us to construct classes of users and their voting sessions.

Each table can be used to derive an object such as an user of a collaboratory, a voting session, etc. However, note that we do not propose a rigorous object-oriented schema. The purpose of referring to objects is to create templates to be used for document construction. Since the schema is more of an object-based setup rather than object-oriented, no reference to polymorphism or message passing is required. As the implementation is Web-based (i.e., the end result is an HTML document), such a modular conception helps increase the efficiency of dynamically constructing HTML documents with embedded objects. Further, such a schema can be extended to corporate Intranet settings where external objects (e.g., information about users from other databases) can be easily brought into the current setting. It is beyond the scope of this article to discuss the complete implementation details.

**************INSERT FIGURE 10 ABOUT HERE**********

5.3.2 EC World: A Collaboratory based on both Intranets and the Internet

The two Collaboratories described earlier are either entirely Internet or Intranet based. However, in a large number of situations, one may require a Collaboratory which combines both the Internet and Intranets. Electronic publishing is one such domain of interest. We created *EC World - A Forum for the 21st Century*[1], an on-line electronic journal exclusively for the field of electronic commerce. EC World is an entirely Web-based journal built on Internet protocols and Relational Database Management Systems (RDBMS) technology. While the primary objective of EC World is to provide a platform for dissemination and creation of knowledge in the area of electronic commerce, from an administrative perspective, we structured it along the model of a networked organization [11].

There are two main facets to EC World, one based on informal collaboration between the relevant actors (readers, authors and reviewers), and the other a completely formal process for article submission and its review. The former takes place within the public Internet domain, while the latter is modeled as a workflow and supported over an Intranet setting to satisfy confidentiality requirements.

The informal collaboration is an example of a special interest group forum where readers post

[1] URL: *http://ecworld.utexas.edu/*

questions, comments, counter-arguments about specific issues in electronic commerce such as EDI, on-line security, payment systems, etc. All users are required to register, while there is a guest account for users who just wish to browse the journal contents.

The formal aspect of EC World is the article submission process, where a prospective author creates a pre-formatted file in HTML and submits it on-line to the journal. During submission the author is required to check areas of interest which closely match the thrust of the article. Based on the author's selection, a subset of reviewers (who are members of the advisory board and have pre-registered themselves as having expertise in certain areas within EC) are automatically chosen and notified of the submission. Upon receipt of notification, the reviewers are expected to connect on-line to a private repository of the journal (which represents an Intranet setting), and to review the article.

For increased efficiency of article handling, the system automatically appends the reviewer's comments and recommendations to the submitted article. Based on the perceived need to avoid a situation where one reviewer gets influenced and biased *a priori* by other reviewer's comments, the system hides the latter from a reviewer's view. In some circumstances, however, the review process may require changes in its workflow, and the technology base of EC World is well equipped to accommodate variations in the workflow components. Once three reviewers have reviewed an article, the system notifies the administrators of EC World of the result, and then the article is either accepted, sent to the author(s) for suggested revisions, or rejected.

6 Conclusion

Dynamically evolving organizational forms and the growing interaction needs of individuals, groups, and organizations necessitate a migration from groupware applications based on proprietary platforms to Internet and Intranet-based Collaboratories. Various Internet technologies such as the Web, USENET, and search engines appear promising for building open Collaboratories, but require major integration amongst themselves as well as with technologies such as databases and scripting to fulfill a wide array of user requirements. In other words, the integrated technology platform offers a more suitable foundation on which Collaboratories can be built. Desirable characteristics of a Collaboratory differ across communities of users, but the choice should be made so as to maximize the net value derived by members of the respective community by selecting a set of complementary characteristics.

We have developed prototypes of various Collaboratories based on the premise that different electronic communities have different functional requirements. For example, due to the adoption of EC, organizational collaboration has more stringent security, authentication, workflow validation, privacy and performance requirements than collaboration in the other two communities. Further, these collaborative activities span multiple units across the organization, making isolated group support systems less valuable than before. By bringing together relational databases and Web technologies in a common orbit, we have created a basic infrastructure for integrating corporate data, business protocols and group systems. Sequel research in this area should focus on how business rules and procedures can be incorporated in the conceptual design, justification and implementation of collaborative systems for EC activities.

References

1. Applegate, L. and Gogan, J. Paving the information superhighway: Introduction to the internet. Technical Report 9-195-202, Harvard Business School, Boston, MA 02163, (August 30 1995).

2. Barua, A., Chellappa, R., and Whinston, A. Creating an MIS collaboratory in cyberspace: Theoretical foundations and an implementation. *Journal of Organizational Computing*, 5, 4 (January 1996), 417–442.

3. Barua, A., Chellappa, R., and Whinston, A. Social computing: Computer supported cooperative work and groupware. In G. Salvendy, ed., *The Handbook of Human Factors and Ergonomics*, New York: John Wiley and Sons, 1996, in press.

4. Barua, A., Kriebel, C., and Mukhopadhyay, T. Information technologies and business value: An analytic and empirical investigation. *Information Systems Research*, 6, 1 (March 1995), 3–23.

5. Barua, A., Lee, C., and Whinston, A. Incentives and computing systems for team based organizations. *Organization Science*, 6, 4 (July 1995), 487–504.

6. Barua, A., Lee, C., and Whinston, A. The calculus of reengineering. *Information Systems Research*, 7, 4 (December 1996), 409–428.

7. Boland, R. and Tenkasi, R. Perspective making and perspective taking in communities of knowing. *Organization Science*, 6, 4 (July-August 1995), 350–372.

8. Chellappa, R., Barua, A., and Whinston, A. Intranets: Looking beyond internal corporate web servers. In R. Kalakota and A. Whinston, eds., *Readings in Electronic Commerce*, Reading, MA: Addison-Wesley, 1996, 311–321.

9. Chen, H., Lynch, K. J., Himler, A. K., and Goodman, S. E. Information management in research collaboration. *International Journal of Man-Machine Studies*, 36, 3 (March 1992), 419–445.

10. Ching, C., Holsapple, C., and Whinston, A. Reputation, learning and coordination in distributed decision making contexts. *Organization Science*, 3, 2 (May 1992), 275–297.

11. Ching, C., Holsapple, C., and Whinston, A. Modeling network organizations: A basis for exploring computer s upport coordination possibilities. *Journal of Organizational Computing*, 3, 3 (1993), 279–300.

12. Daft, R. and Lengel, R. Organizational information requirements, media richness and structural design. *Management Science*, 32, 5 (1986), 554–571.

13. Duhl, J. Integrating objects with relational data: Exploring mapping issues. *Object Magazine*, (August 1996), 71–73.

14. Fulk, J. and DeSanctis, G. Electronic communication and changing organizational forms. *Organization Science*, 6, 4 (July-August 1995), 337–349.

15. Georgakopoulos, D., Hornick, M., and Sheth, A. An overview of workflow management: From process modeling to workflow automation infrastructure. *Distributed and Parallel Databases*, 3 (1995), 119–153.

16. Gupta, S. Introductory remarks - consumer survey of www users. http://www-personal.umich.edu/~sgupta/hermes/survey2/intro2.html, (1996).

17. Hiltz, S. and Turoff, M. *The Network Nation: Human Communication via Computer.* Reading, MA.: Addison-Wesley, (1978).

18. Jarvenpaa, S. and Ives, B. The global network organization of the future: Information management opportunities and challenges. *Journal of Management Information Systems*, 10, 4 (Spring 1994), 25–57.

19. Knoll, K. and Jarvenpaa, S. Learning to work in distributed global teams. In *Proceedings of the 28th Annual Hawaii International Conference in System Sciences*, 4, 1995, 92–101.

20. Kunz, W. and Rittel, H. Issues as elements of information systems. Center for Planning and Development Research, University of California, Berkeley, Working Paper No. 131.

21. Milgrom, P. and Roberts, J. The economics of modern manufacturing: technology, strategy and organization. *American Economic Review*, (June 1990), 511–528.

22. Monge, P. and Fulk, J. Global network organizations. Technical report, International Communication Association, Albuquerque, NM, (May 1995), Paper available from the authors upon request.

23. Neuman, B. Prospero: A tool for organizing Internet resources. *Electronic Networking: Research, Applications and Policy*, 2, 1 (Spring 1992), 1–8.

24. Pickering, J. and King, J. Hardwiring weak ties: Interorganizational computer-mediated communication, occupational communities and organizational change. *Organization Science*, 6, 4 (July-August 1995), 479–486.

25. Pitkow, J. and Recker, M. Using the Web as a survey tool: Results from the second www user survey. *Journal of Computer Networks and ISDN systems*, 27, 6 (April 1995), 809–823.

26. Ramesh, R. and Whinston, A. Claims, arguments, and decisions: Formalism for representation, gaming and coordination. *Information Systems Research*, 5, 3 (1994), 294–325.

27. Rittel, H. A concept for an argumentative planning information systems. Institute for Urban and Regional Development, University of California, Berkeley, Working Paper No. 324.

28. Steinbart, P. and Nath, R. Problems and issues in the management of internal data communications networks: The experiences of American companies. *MIS Quarterly*, 16, 1 (1992), 55–76.

29. Van Maanen, J. and Barley, S. Occupational communities: Culture and control in organizations. *Research in Organizational Behavior*, 6 (1984), 287–365.

30. Walsh, J. and Ungson, G. Organizational memory. *Academy of Management Review*, 16, 1 (1991), 57–91.

31. Westland, C. Self-organizing executive information networks. *Decision Support Systems*, 8 (1992), 41–53.

32. Williamson, O. The economics of organization: The transaction cost approach. *American Journal of Sociology*, 87 (1981), 548–577.

33. Zmud, R., Lind, M., and Young, F. An attribute space for organizational communication channels. *Information Systems Research*, 1, 4 (1990), 440–457.

A APPENDIX: Issues in Web-Database Integration

While connectivity between the Web and database systems can be handled in many ways, the client/server approach offers the maximum flexibility and efficiency in integrating the two technologies. Having decided not to empower the client side (as discussed in the design philosophy section) we focused on a server-centric approach, which can be implemented in at least three ways: Server Application Programming Interfaces (API), proprietary servers, and CGI Scripting. These options are briefly discussed below.

1. **Server APIs:** Many commercial Web servers offer proprietary APIs that can be utilized to build special programs to communicate with a database server. Examples include NSAPI from Netscape Communications for the Netscape Web Server and ISAPI from Microsoft Corporation for the Internet Information Server. These APIs help in building extensions to the Web server itself without making any changes to the HTTP protocol. These extensions

can be dynamically linked to the Web server process, and can therefore obviate the need to spawn separate processes such as in a CGI environment.

Although these APIs offer greater control over managing system resources, they are restricted to vendor specific Web servers. For example, a gateway constructed using NSAPI cannot be used in conjunction with a NCSA Web server.

2. **Proprietary Servers:** Proprietary Web servers differentiate themselves by offering a wide range of functionality including the ability to handle Structured Query Language (SQL) statements. These servers can be considered as extended database drivers that recognize HTTP. For example, Oracle's own Web server greatly extends the HTTP architecture, and supports many programming interfaces that can be used to connect directly to its database engine. It also integrates many other features such as security and authentication, and offers extensions to Java, PL/SQL, etc.

   The architecture of these Web servers varies widely. In spite of their rich functionality, many of these Web servers are tightly coupled with their respective database environments. For example, the Oracle Web server cannot be fully exploited with a Sybase database engine. This may be a serious limitation for a gateway in a heterogeneous multi-vendor environment.

3. **CGI Scripting:** CGI is a way of redirecting HTTP requests to programs that reside outside a Web server environment, such as a database server. This approach does not require any changes to the server or the client, and was deployed in developing our database-enriched Collaboratories. In this method, an HTML form is used to input data, which is then is then passed on to a CGI script. The script sends the data to a database server in an appropriate manner (discussed below), and the response from the database is received by the script and formatted in HTML or other MIME types that a client can handle.

   The CGI script itself can be divided into three parts: the parser, the database connector and the formatter. The parser breaks down the HTML form inputs into variables that can be passed to SQL statements. The database connector involves some considerations, where a process pipe needs to be created to log on to the database. One approach is to embed SQL statements in the script and to redirect the output of this script to a command-line database client.

For example, SQLPLUS is a command-line interface to a ORACLE DBMS. Normally this is used to write SQL statements in an interactive fashion, and the data can be viewed on the screen. However, the same approach could be used within a scripting language such as PERL as a sub-routine to simulate interaction. The following pseudo-code highlights this approach.

```
query = SELECT something FROM sometable;

sub ora_query{
        local($database, $query) = @_;
<<'eoc';
/exhome/orahome/bin/sqlplus -s $database <<eof
set echo off
set flush off
set embedded on
set verify off
set feedback off
set heading off
set pagesize 0
set wrap on
$query
exit
eof
eoc
}
```

Although this technique achieves basic connectivity, problems may arise in redirecting output from the constructed query. Instead of invoking command-line utilities, it is also possible to construct gateways by using generic calls that are supported by the database itself. For example, ORACLE provides extensions to many programming languages such as C, FORTRAN, ADA, etc., that can be used to construct gateways. Commonly known as Oracle Call Interface (OCI), language extensions such as ProC for the C language provide libraries that can establish database connectivity, and a simple C program can be constructed to redirect the parsed values from the HTML form to the database through embedded SQL statements. Similarly other languages such as PERL have extensions known as OraPerl that help attain the same functionality. The following example illustrates how a database call and subsequent select statements can be invoked directly from within a CGI program.

```
$csr=&ora_open($lda,"SELECT something FROM sometable WHERE
articleID=$articleID") || print"Error ";
($decision)=&ora_fetch($csr);
```

This method also suffers from many inadequacies. Since all information provided and received through HTTP need to be done in one pass, a new process is spawned for each request. Thus, regardless of the amount of data requested, a new database connection is initiated each time and consumes a large amount of system resources. Further, it also suffers from the lack of flexibility in that the client cannot construct a SQL statement but can only invoke pre-designed embedded procedures. However, the key advantages of this approach are its simplicity and open nature.

B Appendix: Tables and Figures

	Electronic Communities		
Community-Specific Characteristic	**Informal Groups**	**SIG**	**Organizations**
Intellectual Organization of Information and Interaction	√	√ √	√ √ √
Validity/Reliability	√	√ √	√ √ √
Workflow Support		√	√ √ √
Common Infrastructure	Information Access		
	Interaction Richness		
	Information Scalability		

Table 1: Desirable characteristics of collaborative systems

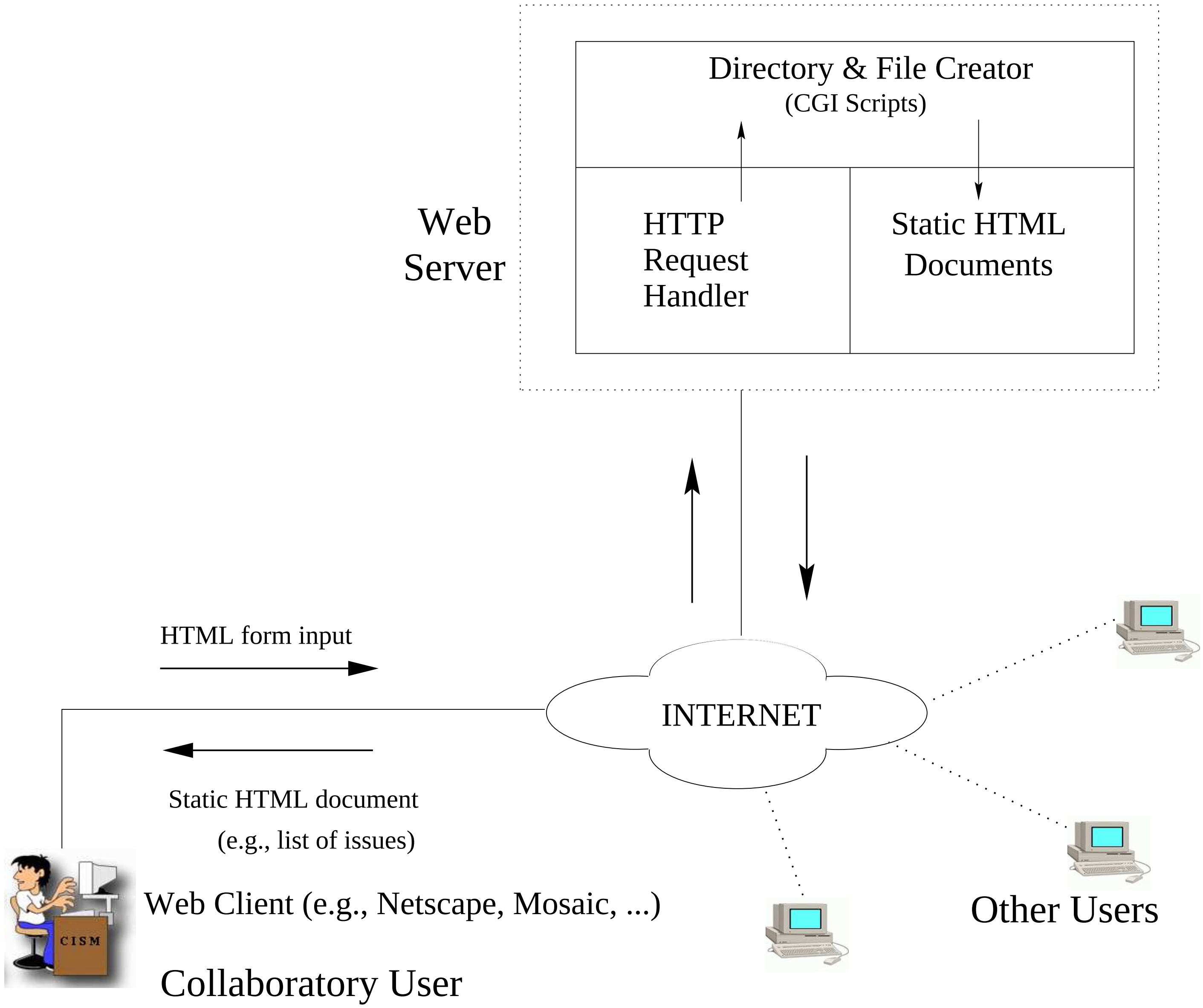

Figure 1: Flat-file based Collaboratory

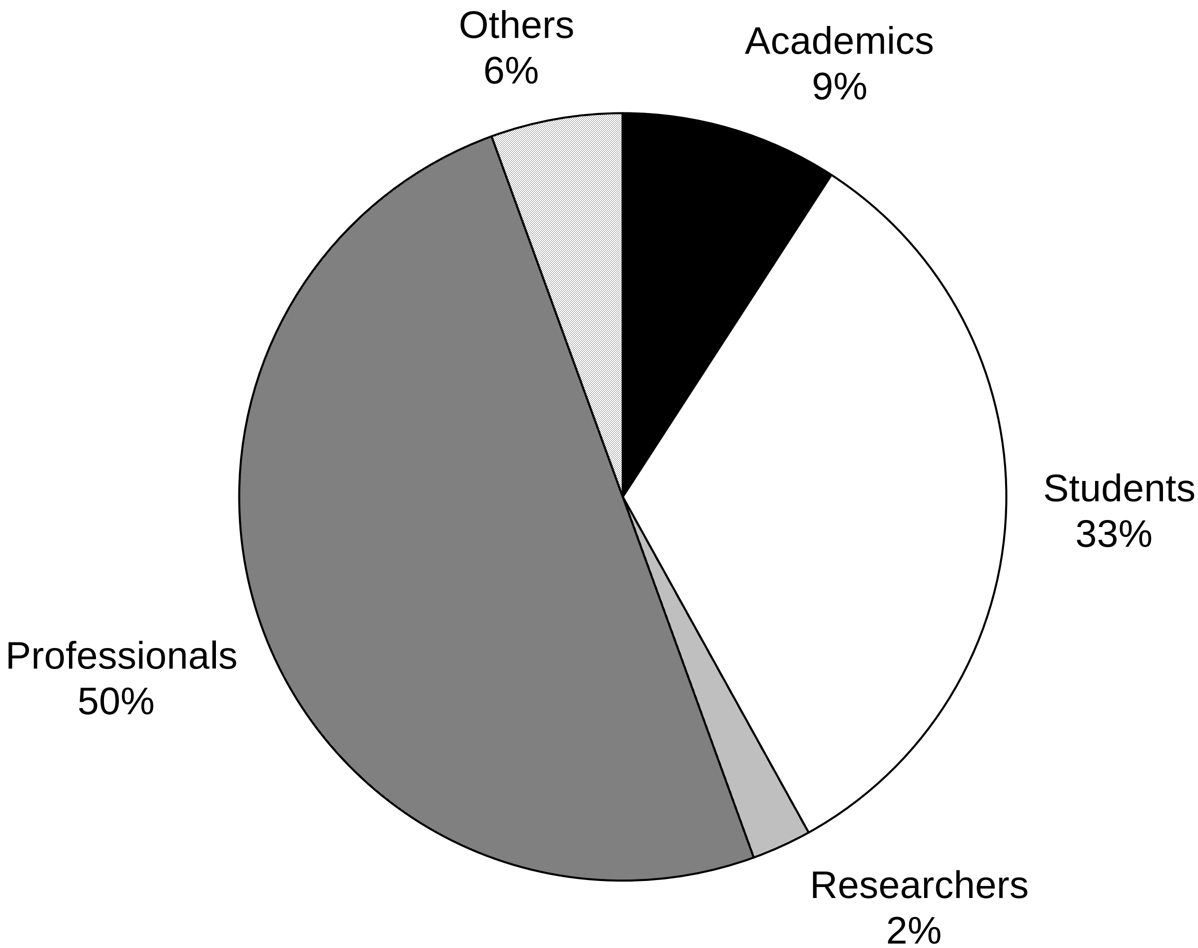

Figure 2: Respondents' occupations

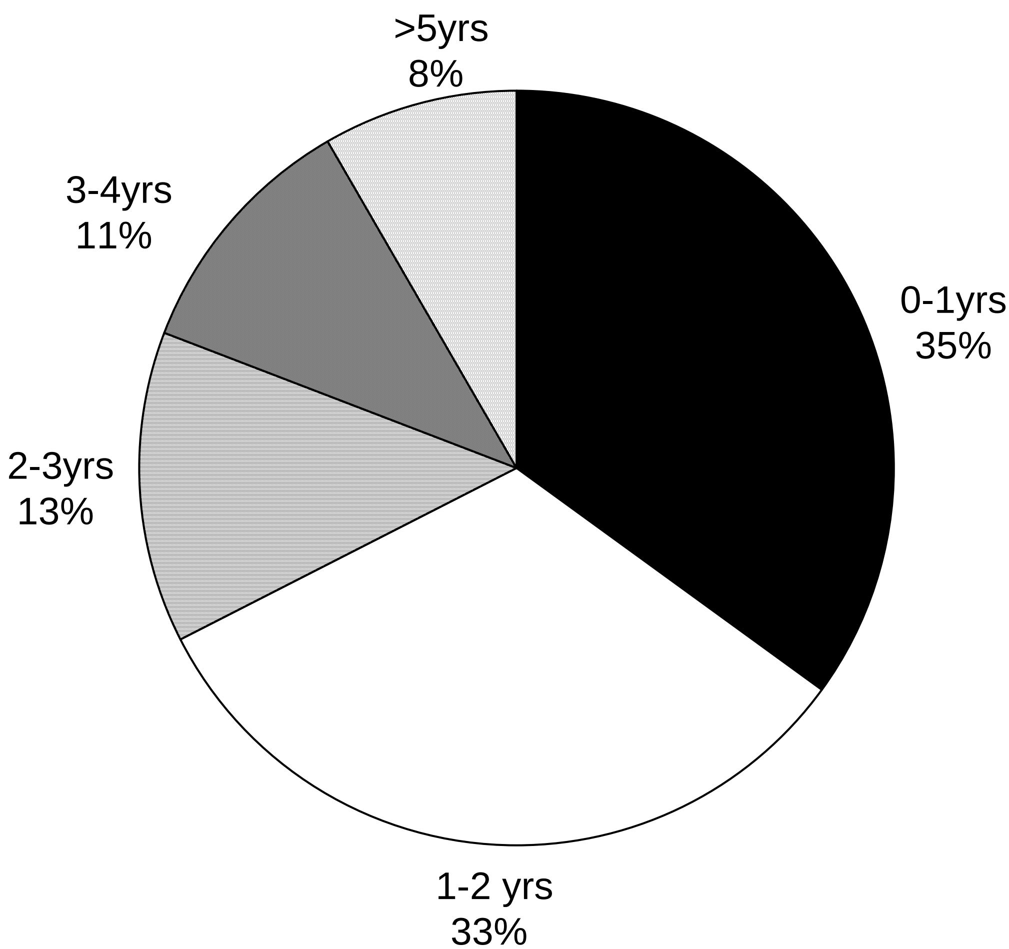

Figure 3: Internet experience

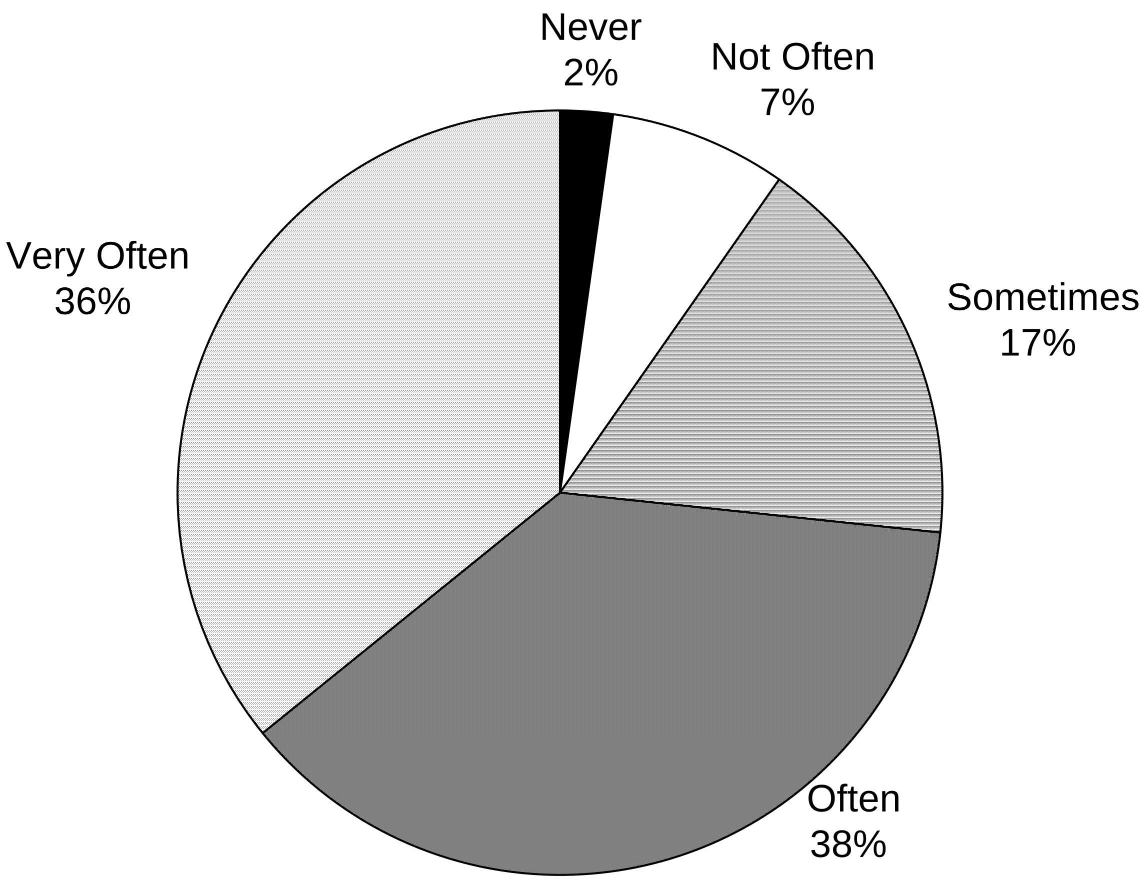

Figure 4: Work-related Internet usage

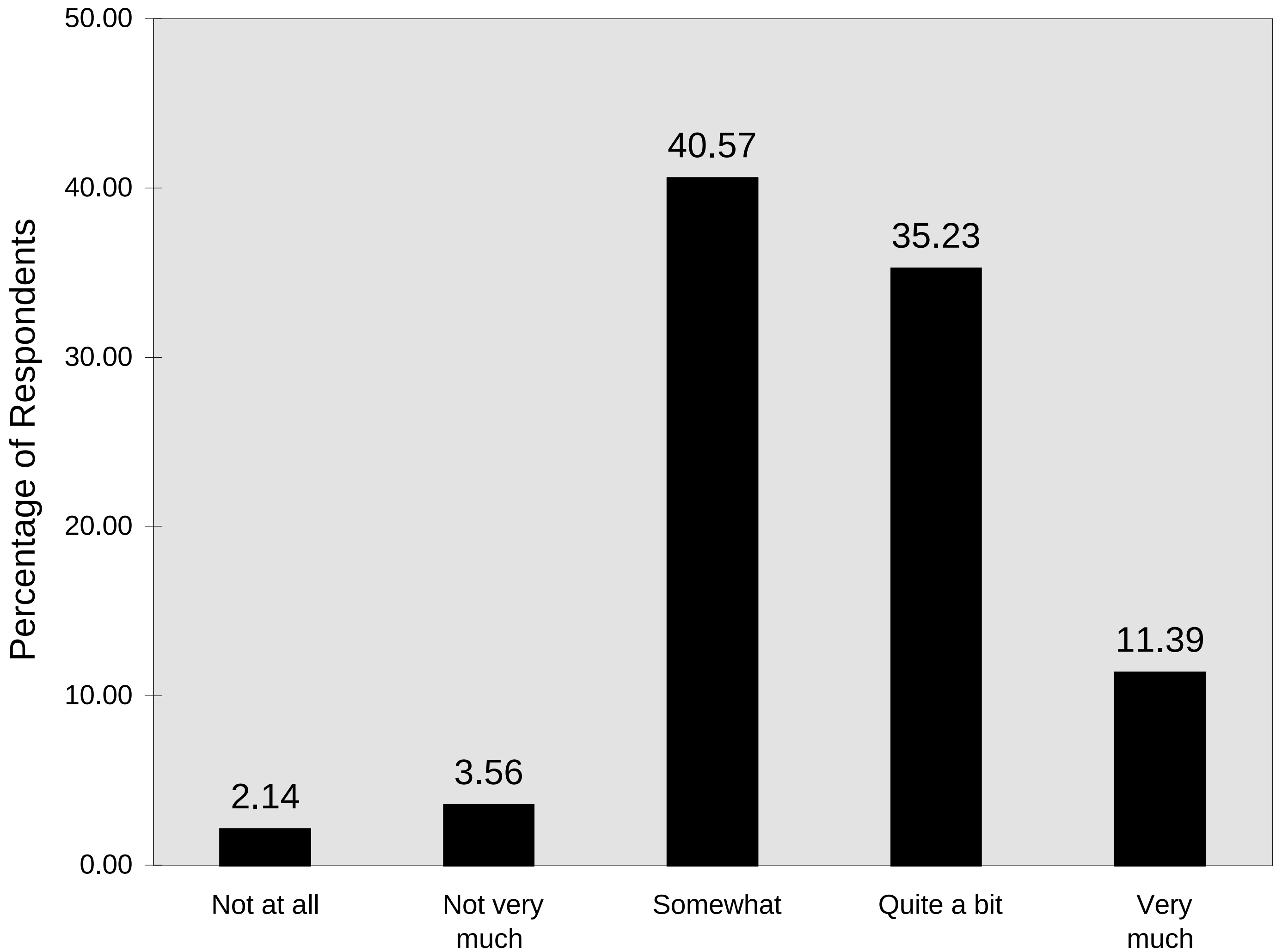

Figure 5: The MIS Collaboratory's usefulness as a platform for special interest group interactions

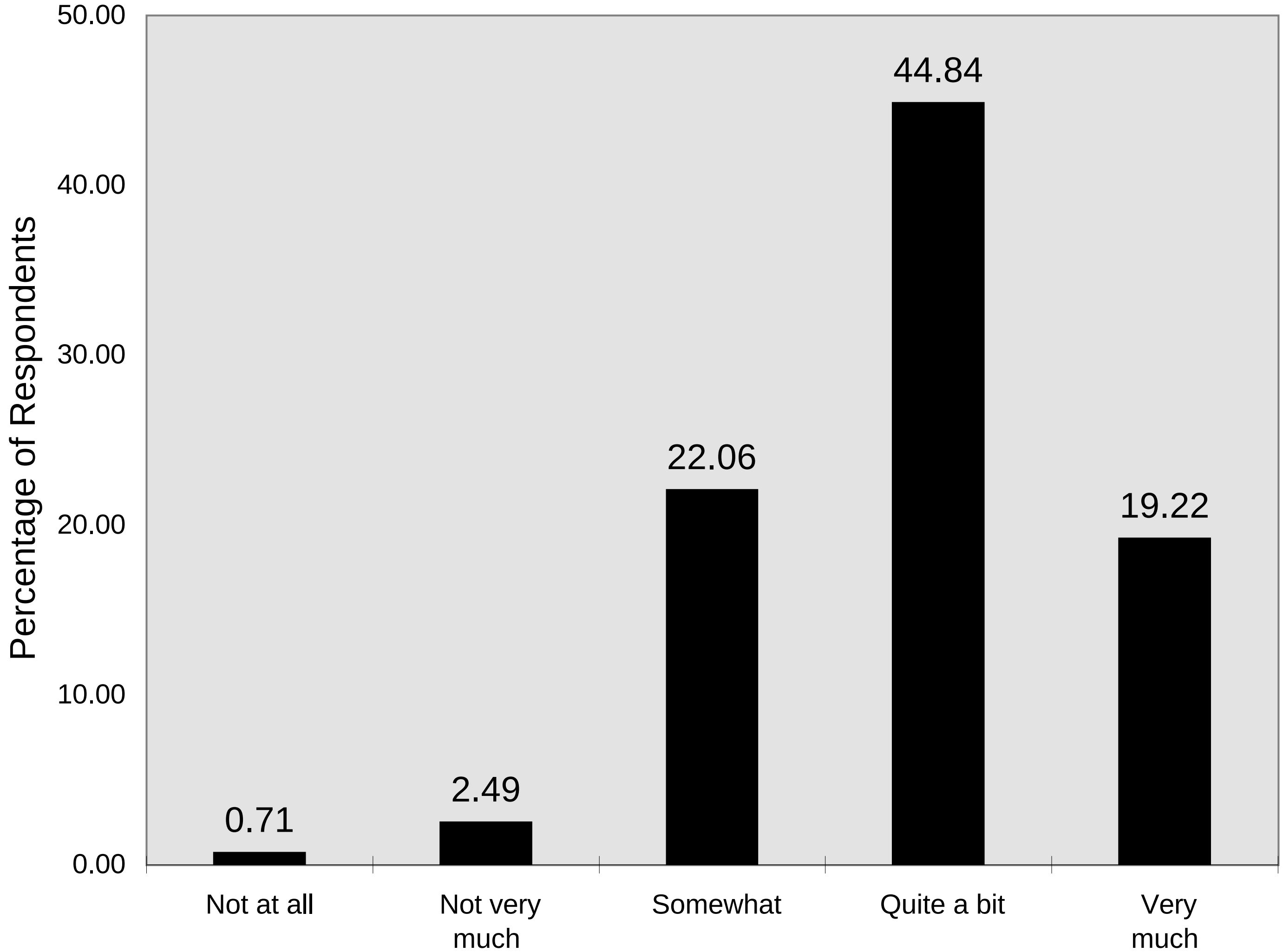

Figure 6: The MIS Collaboratory's usefulness relative to newsgroups

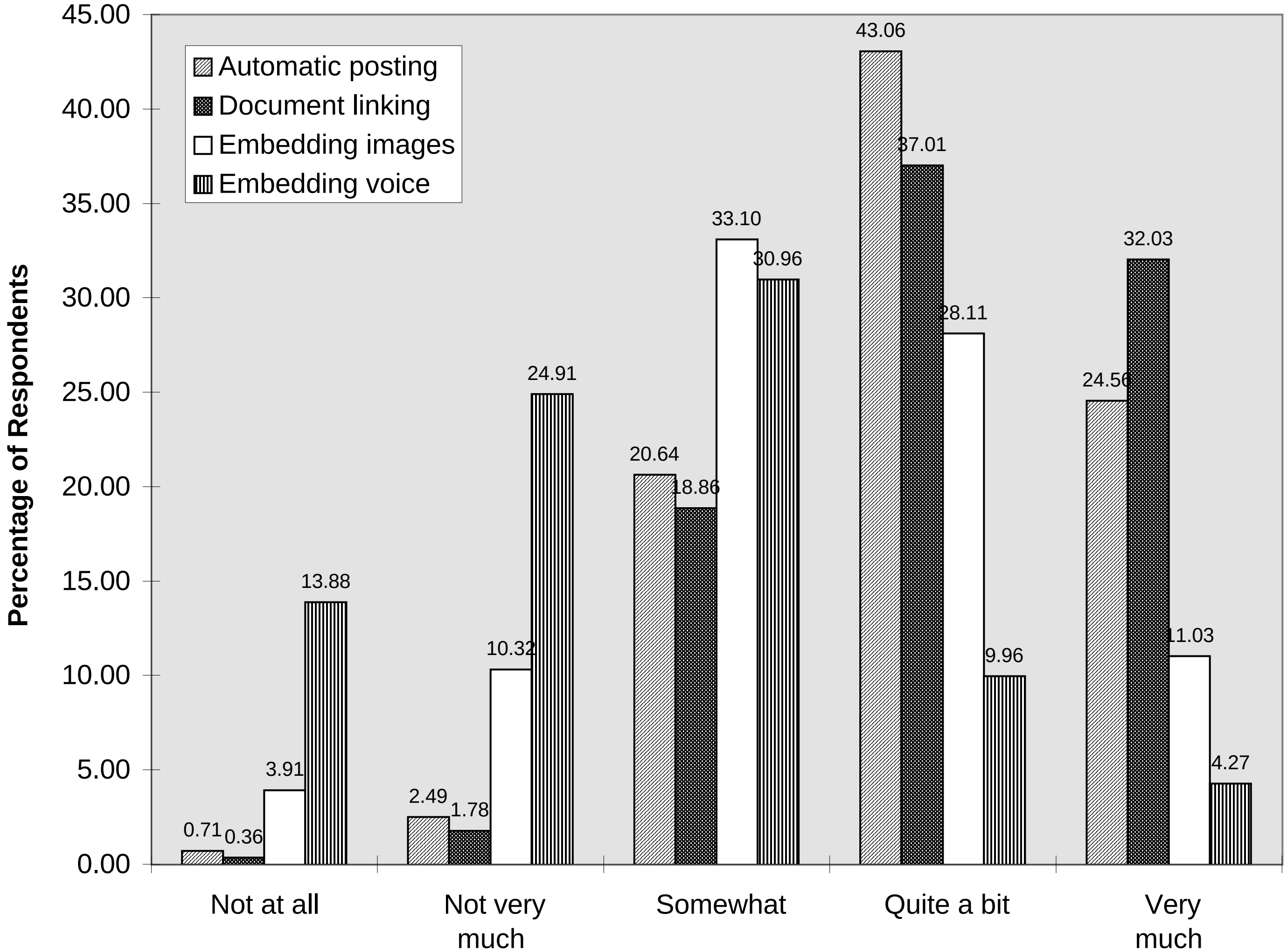

Figure 7: Assessment of the MIS Collaboratory's features

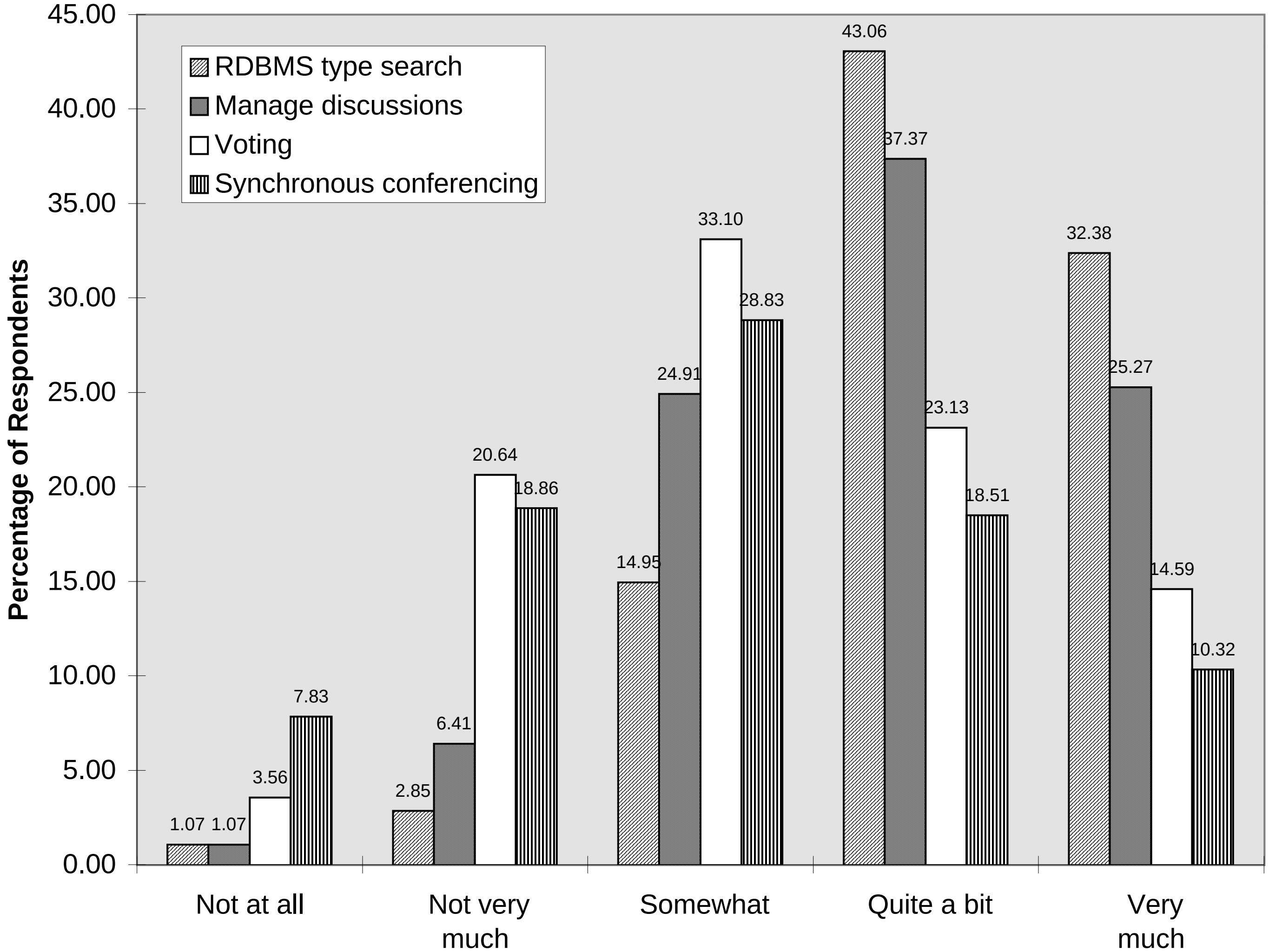

Figure 8: Perceived importance of planned features

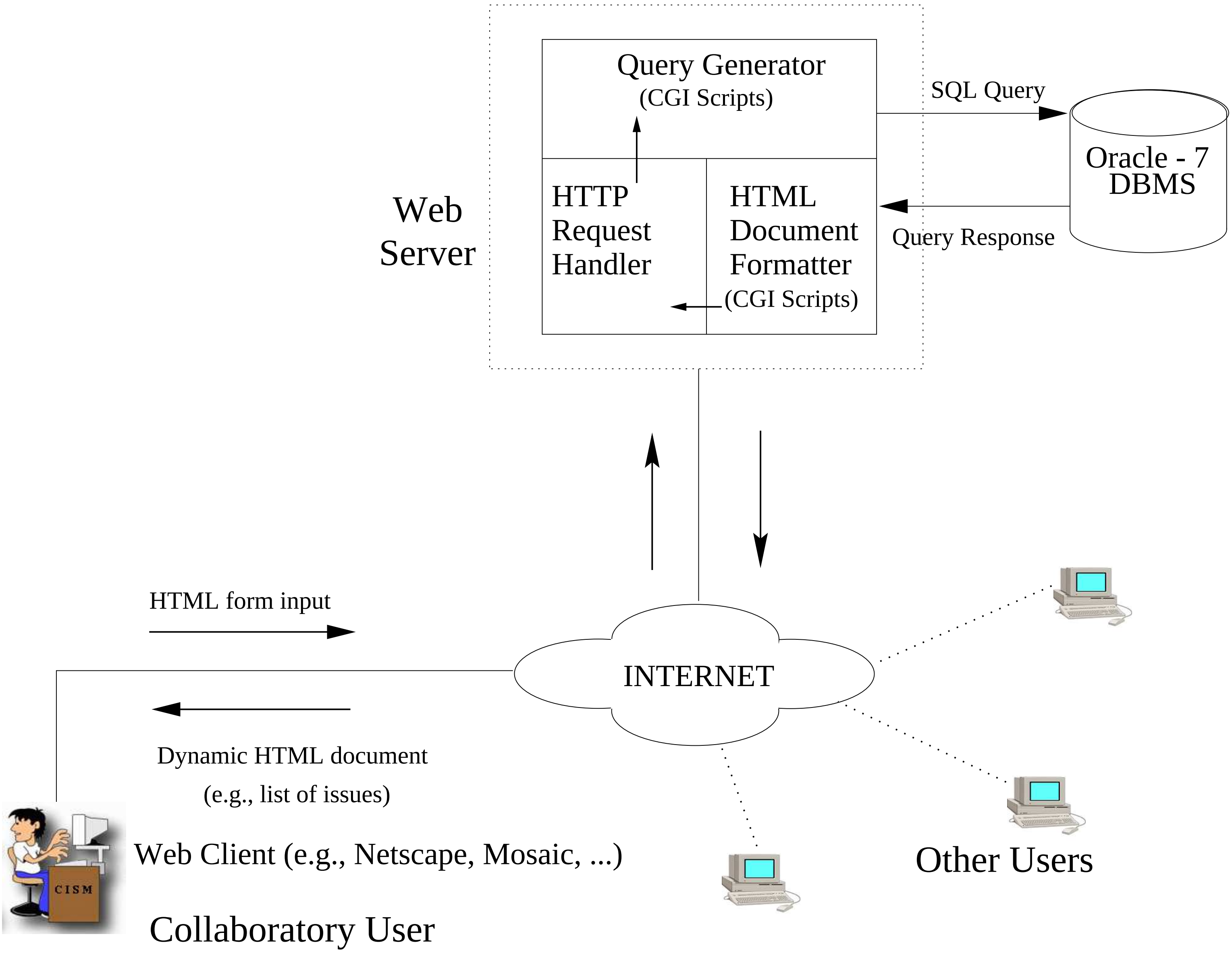

Figure 9: Database-enriched Collaboratory

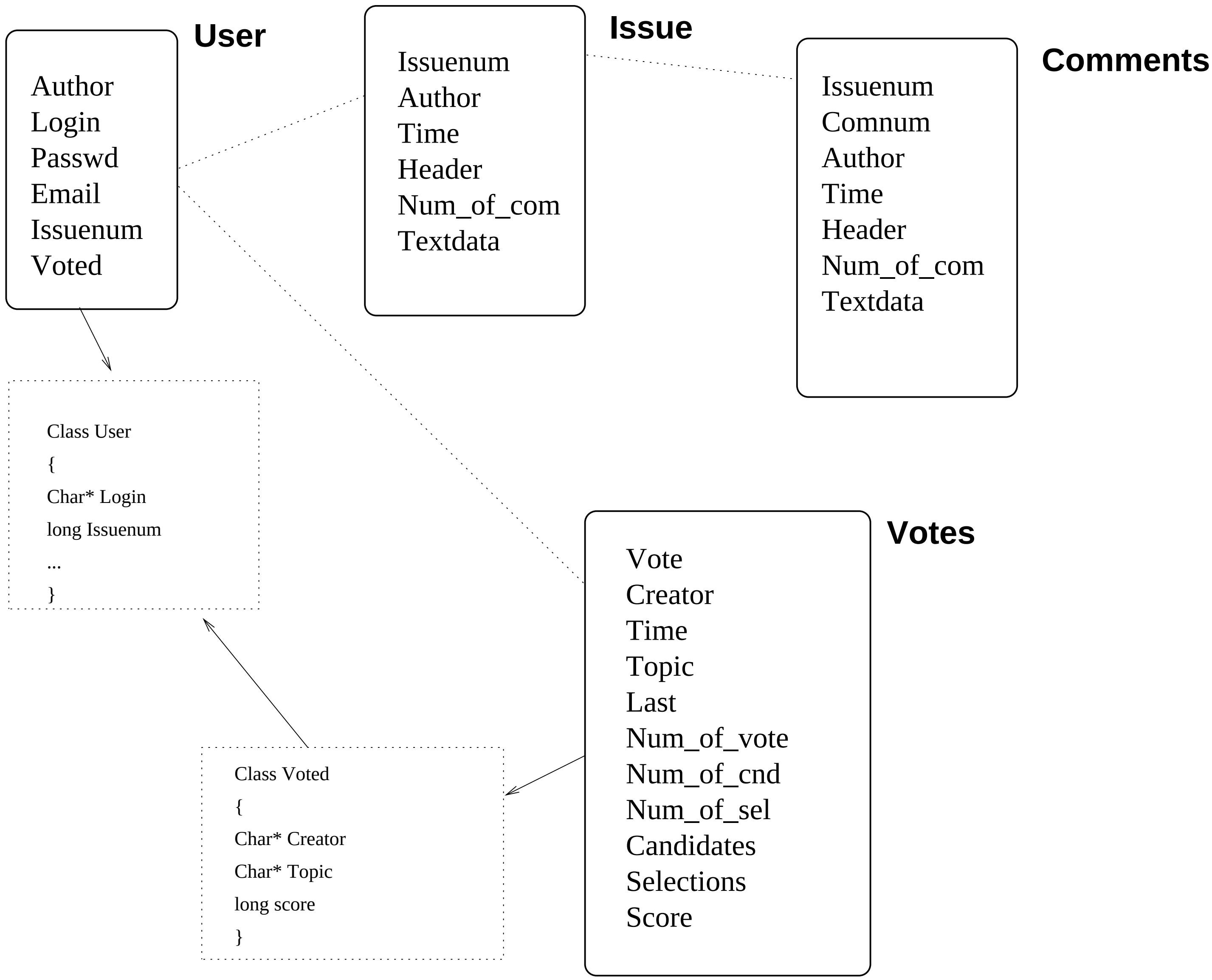

Figure 10: A partial schema of the database-enriched Collaboratory