Published in final edited form as:
Int J Med Inform. 2009 January ; 78(1): 10–21. doi:10.1016/j.ijmedinf.2008.06.011.

Incorporating collaboratory concepts into informatics in support of translational interdisciplinary biomedical research

E. Sally Lee[a,*], **David W. McDonald**[d], **Nicholas Anderson**[a], and **Peter Tarczy-Hornoch**[a,b,c]

[a] Division of Biomedical and Health Informatics, Department of Medical Education and Biomedical Informatics, University of Washington, Seattle, WA, USA

[b] Department of Pediatrics, University of Washington, Seattle, WA, USA

[c] Department of Computer Science and Engineering, University of Washington, Seattle, WA, USA

[d] The Information School, University of Washington, Seattle, WA, USA

Abstract

Due to its complex nature, modern biomedical research has become increasingly interdisciplinary and collaborative in nature. Although a necessity, interdisciplinary biomedical collaboration is difficult. There is, however, a growing body of literature on the study and fostering of collaboration in fields such as computer supported cooperative work (CSCW) and information science (IS). These studies of collaboration provide insight into how to potentially alleviate the difficulties of interdisciplinary collaborative research. We, therefore, undertook a cross cutting study of science and engineering collaboratories to identify emergent themes. We review many relevant collaboratory concepts: (a) general collaboratory concepts across many domains: communication, common workspace and coordination, and data sharing and management, (b) specific collaboratory concepts of particular biomedical relevance: data integration and analysis, security structure, metadata and data provenance, and interoperability and data standards, (c) environmental factors that support collaboratories: administrative and management structure, technical support, and available funding as critical environmental factors, and (d) future considerations for biomedical collaboration: appropriate training and long-term planning. In our opinion, the collaboratory concepts we discuss can guide planning and design of future collaborative infrastructure by biomedical informatics researchers to alleviate some of the difficulties of interdisciplinary biomedical collaboration.

Keywords

Collaboration; Biomedical informatics; Computer supported collaborative work; Collaboratories; Social and technical issues; Bioinformatics

1. Introduction

The modern biomedical research community faces many exciting and challenging research problems. The complex nature of modern research questions has led to a realization in the research community that a single lab or a single discipline often can no longer provide all the necessary expertise and resources to solve these questions [1]. Thus, biomedical research has become increasingly interdisciplinary and collaborative in nature [1–4]. The National Institute of Health (NIH) recognized the need for interdisciplinary collaboration in its roadmap for

*Corresponding author at: Box 357240, University of Washington, Seattle, WA 98195-7420, USA. Tel.: +1 206 616 0369; fax: +1 206 543 3461. E-mail address: sallylee@u.washington.edu (E.S. Lee).

modern biomedical research in the 21st century [5]. An element of this recognition of the need for interdisciplinary collaboration is the growing recognition of biomedical informatics as an important part of the foundation to support interdisciplinary research. A concrete example of the crucial function biomedical informatics will play in future interdisciplinary collaboration is described in the executive summary of the National Meeting on Enhancing the Discipline of Clinical and Translational Sciences [6].

As necessary and valuable as this is recognized to be, the literature shows that interdisciplinary scientific collaboration is not easily established or maintained for several reasons. First, the researchers involved typically possess a wide range of expertise from multiple disciplines with different cultures and social norms with no common ground to draw upon to seamlessly communicate [1]. Second, biomedical research in particular is highly competitive and many researchers involved are unwilling to share and trust [7–9]. Third, many scientists are reluctant to trust unfamiliar tools which often results in lack of adoption of core technologies [10]. Fourth, physical proximity is known to be important to scientific collaboration due to the necessity of informal communications in fostering collaborative environment, yet interdisciplinary biomedical researchers often lack proximity with their collaborators [11,12]. Fifth, there is a general lack of a common infrastructure connecting all disparate systems and workflows due to disparate locations, disciplines, and lack of adoption [13,14].

There is a growing body of literature in fields such as computer supported cooperative work (CSCW) and information science (IS) on fostering collaboration that provides insight into how to potentially alleviate the difficulties of interdisciplinary collaborative research [15–39]. One prominent concept that has been heavily researched in the fields of CSCW and IS, is the “collaboratory” [40]. From collaboratory research emerged many social and technological factors associated with collaboration that are useful to consider for facilitating interdisciplinary biomedical collaboration. For biomedical informatics to support biomedical research as it becomes more collaborative, the necessary social and technical aspects of support will likely be similar to those identified in collaboratory research in the broader sense described previously. Thus our opinion is that understanding and learning from the existing collaboratory research is important.

We undertook this literature review for two key reasons. First, none of the existing collaborative research has addressed issues of collaboratories specifically from the biomedical, clinical, and translational research point of view. This raises questions about where the important overlaps exist and where there are gaps that must be addressed when implementing collaboratories for biomedical, clinical, and translational research. Second, the existing collaboratory literature has not thus far included a comprehensive review article that includes the recurring elements across infrastructure, social and environmental issues that are critical to collaboratory. This review brings together the wide and diverse range of collaboratory results, some of which might be overlooked by others in their rush to develop novel technologies. The goal of this review is to identify concepts in the collaboratory research literature that can inform biomedical informatics to better support interdisciplinary biomedical research.

In this paper, we will provide a systematic definition of a collaboratory and a thorough review of collaboratory concepts. We show how general collaboratory knowledge can be applied to benefit collaborative biomedical research. Furthermore, we identify more specific aspects of collaboratories that will be particularly applicable and needed for informatics systems in support of collaboration in the biomedical field. In order to achieve these goals, we first identify general concepts that are necessary in all the collaboratories; then proceed to describe specific collaboratory concepts that in our interpretation are desirable in biomedical collaboration.

2. Background

2.1. History of collaboratories

Collaboratories are “laboratories without walls” where researchers can perform their research independent of time and location. Here, they can interact with colleagues, access instrumentation and information, share data and computational resources [40,41]. Although the need for remote collaboration has been recognized since the introduction of the collaboratory concept in 1989, no infrastructure existed to support such research. Hence, early collaboratories were initially focused on developing technology to support remote collaboration such as tools for communication and shared access to instrumentation. For example, the Worm Community System (WCS) was one of the first collaboratories to form that provided a collaborative scientific research system for geographically dispersed biologists [40,42,43].

Many early collaboratories were formed in engineering disciplines since remote access to expensive equipment or instrumentation drove the creation of these engineering collaborations. Two notable examples of this type of collaboratory include the Upper Atmospheric Research Collaboratory (UARC) [40,44] and the Environmental Molecular Sciences Laboratory (EMSL) [40,45,46]. Many of the collaboratories that followed, such as Biological Collaborative Environment (Bio-CoRE) [47,48] and the Biological Sciences Collaboratory (BSC) focused on building systems to support data access, management, and analysis [49, 50].

The term “collaboratory” historically described simple collaborations focused on the use of technology to access remote instrument or data – the term is no longer widely used in this sense today. Instead, the concept of collaboratories has evolved and expanded to large collaborations and consortiums spanning multiple institutions, covering multiple domains of science, and leveraging various collaboratory technologies. The collaboratories are now called by other names, such as consortium, eScience, Grid, center, core and network. The nature of collaboratory study has also changed. Initially, tool building was necessary to enable remote collaboration; however, as the collaboratory technologies matured, the focus of collaboratories shifted toward the support of both social and technical processes of scientific research as whole.

2.2. Theories on collaboratories

There is no all-encompassing theoretical framework that has been used to describe collaboratories and the related informatics issues. Collaboratories are complex socio-technical systems that belie a one-solution-fits-all approach. Indeed, most collaboratories are custom designed, each tailored to meet certain needs. Yet due to the problems and high costs of one-off developments, nearly two decades of collaboratory research have sought to identify a common technical and social environment that would facilitate reuse of existing infrastructure while maximizing the adoption and use. The problems in collaboratory research are therefore broken into pieces that can be addressed by the different disciplinary approaches and theories rather than the overarching ‘collaboratoryness’.

Collaboratory research has addressed these different areas:

- *Design* – What is the best way to engage the design of a new collaboratory?
- *Features* – What are the most important technical features and how do users interact with or operate those features?
- *Sharing* – How and when are data or physical equipment shared?

- *Communication* – How do participants in the collaboration interact with each other? Should the interaction be through the collaboratory explicitly or through other means?
- *Coordination* – What mechanisms allow participants in a collaboratory to effectively coordinate their individual activities?
- *Social Factors* – How do social relationships, both existing and new, influence the perception of the system and the collaborative effort?
- *Environment* – How does the organizational and institutional frame in which the collaboratory and the participants sit influence the technology and the collaborations?

While there is no single theory that cuts across the study of collaboratories, there are several methodological approaches that have been useful in conceptualizing technical and social issues. For example, many of the collaboratory studies involved researchers who engaged the design of collaboratory infrastructure using Participatory Design (PD) [51–54] and User Centered Design (UCD) methods [55]. By the time collaboratory research had begun, the "strong" design view that theoretical models of collaboration should guide design had largely been abandoned. Prior attempts at theory driven designs, like that in The Coordinator or Action Workflow, two systems that implemented models of human collaborative activity based on Speech Act Theory, have met with much user resistance and failure [56]. Adopting PD and UCD methods allowed early collaboratories to begin addressing domain specific issues of collaboration. These approaches to design cut across the majority of the collaboratory literature and are still in use by many collaboratory developers.

There are few other methodological or theoretical approaches which cross the majority of the research on collaboratories. This does not mean there is no commonality, just that the commonality is driven more by the discipline of the researcher and the specific problem under investigation. Only within last few years, the dearth of common theoretical frameworks for designing and studying collaboratories lead to the elaboration of TORC (Theory of Remote Collaboration). The TORC is a descriptive theory that is based on a set of evaluation studies and a survey about the critical aspects of remote collaboration [57].

3. Method: approach to literature review

For this review, we surveyed the literature in a variety of relevant disciplines that study and support collaborations. We specifically surveyed the literature in Computer Science, Information Science, Biomedicine and Social Sciences. The focus of the search was in technical, as well as social issues related to collaboration that look explicitly at collaboratory literature as well as related literature such as CSCW. The primary search sites to access literature were Google Scholar [58] and PubMed [59]. We included in our search to other relevant literature indices and databases such as: Library Literature and Information Science, Library and Information Science Abstracts (LISA), ISI web of Knowledge (Social Sciences Citation Index), Association for Computing Machines (ACM) and the Institute for Electrical and Electronics Engineers (IEEE). We used published articles, books, conference proceedings, and grey literature such as dissertations and websites. The initial keywords searched were: collaboratory, collaboratories, and biomedical collaboration. Subsequent search was performed through citation and reverse citation search as well as through discussions with colleagues in the area of collaborative research.

Due to the heterogeneous nature of literature that covers collaboratories, the research cited in this review includes a wide range of publications from different disciplinary perspectives. The citations cover formative studies designed to collect requirements for new collaboratories, participatory design studies, systems design and deployment studies, as well as qualitative summative evaluations of collaboratories in use. The approach taken by these studies varies from descriptions of the systems to analyses of the systems without a formal methodology to qualitative analyses of themes across collections of systems. We restricted our review only to the science and engineering research collaboratories. Although other types of collaboratories such as educational or film collaboratories exist, the challenges within the science disciplines are unique.

For each selected area of literature, we carefully read and developed a list of concepts mentioned. After this list of concepts was compiled, we sorted and categorized individual concepts so that they can be described in a simpler way (i.e. grouped into the most frequently mentioned concepts or less frequently mentioned but relevant concepts). In each of these sorting and categorization processes, the list of concepts and the categories were validated with co-investigator in the field of CSCW, and then cross-validated with another co-investigator who is in the field of biomedical informatics. Because of the heterogeneous nature of the articles being reviewed, much valuable literature is not captured in the existing frameworks but instead in the primary literature describing individual systems. We elected to take a qualitative approach to the literature characterizing and coding the themes addressed in each article to identify critical results that cross methodological approaches and disciplinary boundaries. We took this approach rather than the more traditional approach of reviewing serially the existing frameworks from a biomedical informatics perspective because we felt that qualitatively categorizing the emergent and recurring themes would present a more synthetic overall view of the existing literature.

4. Results: review of collaboratory concepts

To provide systematic definition and a thorough review of collaboratory concepts, we have compiled a list of collaboratory concepts (see Table 1). All the collaboratory concepts we have observed in the literature were organized into: (a) general concepts relevant for any collaboration, (b) specific concepts relevant for biomedical collaboration, (c) environmental factors that support collaboration, and (d) other collaboratory concepts. In next few sections, we will illustrate different aspects of all of these concepts. For each concept in Table 1, we provide a brief definition of each of these collaboratory concepts referenced in the literature.

4.1. Review of general collaboratory concepts relevant to any research domain

We identified three essential general collaboratory concepts that are both technological and social in nature. The literature suggests that interdisciplinary collaborations cannot function without these collaboratory concepts incorporated into their structure. They are: (1) communication, (2) common workspace and coordination, and (3) data sharing and management.

4.1.1. Communication—Communication and the exchange of ideas and thoughts is a foundational component of collaborative scientific work. In scientific collaboration, all those involved in collaborative research formally and informally work together, discuss and analyze their research, and share ideas. These formal and informal communications foster collaborative relationship by improving working relationships and maintaining shared knowledge [12]. The importance of informal communication in scientific collaboration, such as chance meetings in the hallways, is well documented [11,12,60,61]. Furthermore, occasional face-to-face meetings are found to be crucial for effective use of communications tools at distance [62]. It is important to note that one reason remote collaborations often fail is partly due to the lack of these informal communications [11].

In interdisciplinary scientific collaboration, the process of communication is tightly coupled with communication technology because researchers are often dispersed in remote locations. Communication technology may be synchronous or asynchronous. Through synchronous tools such as phone, videoconference, or chat systems, scientists can talk to each other in real time. Real time communication is important because certain methods of communication such as brainstorming cannot otherwise easily occur. Asynchronous tools such as e-mail or discussion lists enable scientists to communicate without having to set up a specific meeting time; thus, the researchers are able to work around their busy schedule by using asynchronous communication mechanisms. Communication technology can facilitate both planned and unplanned interactions to promote collaboration and awareness of other's research [1,12,60, 63,64].

4.1.2. Common workspace and coordination—A common workspace refers to a place where researchers can work together [65]. It is needed for work coordination, communication, and informal information and knowledge transfer [66]. Historically, scientific research typically happened in a single lab; thus having a common space was not an issue. With advent of interdisciplinary collaboration brought on by complex scientific problems and funding initiatives, having a common space is often no longer possible [2]. Interdisciplinary collaboration frequently involves scientists in multiple institutes and labs which are geographically dispersed. However, scientific research cannot succeed without some type of common workspace where researchers can work together. Lack of common physical space is especially problematic since as proximity decreases, communication and coordination decrease, and collaboration has an increased chance of failing [12].

In order to alleviate the problem of a lack of shared space, common workspace and coordination technologies have been developed to give scientists a simulated co-located space. A virtual workspace is where all the collaboration related information (e.g., data, communication or software) are collected and structured, and can be accessed by the researchers involved to coordinate and work together [67]. Through this shared space technology, researchers can come together, share data, analyze findings, and have research discussions. In case of collaboration, such technology is typically a system accessible by all the researchers involved in the research collaboration such as Basic Support for Cooperative Work (BSCW) [67,68] or a wiki [69]. BSCW is "a 'shared workspace' system which supports document upload, event notification, group management and much more" [70]. A Wiki is "a collaboratively created and iteratively improved set of web pages, together with the software that manages the web pages" [69]. Although it is not a physically co-located workspace, common workspace and coordination technology can facilitate a simulated virtual common space. It is important to note that such technology is known to reduce negative impact of remote collaboration [2].

4.1.3. Data sharing and management—In collaborative research, data are often generated from multiple individual groups. Each individual group must have an ability to effectively manage data generated as well as the collaboration as whole. Data management technologies help effectively handle data created during research both in smaller group settings and in collaboration as whole. Furthermore, data generated in collaborative research are often shared and analyzed by all those involved in the collaboration. Hence, data sharing and management go hand in hand in collaborative research. Data sharing is beneficial for several reasons: (1) a researcher's findings can be validated by peers, (2) new analyses can be performed on existing data, (3) data can serve as the basis for new research, and (4) it can prevent unnecessary duplication of effort [71,72]. Further more, new knowledge and insight can be obtained from collaborative data sharing that might not be discovered examining individual data. For these reasons, funding agencies such as NIH and NSF have been increasingly requiring data sharing as part of their funding initiatives [6,73,74].

Data sharing often involves technology that enables scientists to exchange data. Such technology could be as simple as spreadsheets which could be sent over an e-mail, or more complicated systems such as a common database repository where all researchers can submit their data. Even if there is a mechanism to share data, sharing would not be possible without trust and willingness to share. Although data sharing is perceived as beneficial, it is hard for scientists to share data due to the competitive nature of scientific research and the historical culture of not sharing [8]. Researchers often do not trust others involved in collaboration, thinking that when data is shared, someone might steal their findings [57,72], a mistrust that stems from highly competitive nature of biomedical research [8,9,72]. When implementing data sharing technology, such social barriers should be taken into account.

4.2. Review of collaboratory concepts relevant for collaborative biomedical research

We have identified four specific concepts in collaboratory research that are particularly relevant to biomedical collaboration. Although they are not concepts applicable to all collaborations, our interpretation is that they are essential collaboratory concepts for collaborations in the biomedical arena. They are: (1) data integration and analysis, (2) security structure, (3) metadata and data provenance, and (4) interoperability and data standards.

4.2.1. Data integration and analysis—Modern biomedical research increasingly generates large amounts of data; however, data without analysis is meaningless. For example, an image generated by a microarray experiment is useless without relevant statistical analysis to indicate what part of that image is significant [75]. In collaborative research, the researchers often collectively analyze data as well as review individually analyzed data. A collaborative analysis can be defined as “an interactive process of brainstorming where researchers share their individual interpretations, understanding, and insights, which build upon one another to form cogent findings and conclusions” [50].

Such collaborative data analysis often involves sophisticated statistical or analytic tools as well as a mechanism for the researchers to tie all the analysis processes together [50]. Such integrated analysis systems enable scientists to view all the data in one place, share analysis tools, and capture notes, working ideas, and interpretations. One important function of this type of analysis system is data integration. Without integrating data generated by all the researchers involved in the collaboration, data cannot be analyzed or reviewed collaboratively. Data integration is also a concept widely known and is actively being researched in the biomedical field due to advent of data explosion challenges [76,77].

4.2.2. Security structure—Biomedical research often involves highly sensitive data such as individual medical records or genetic information. In order to protect these sensitive data, a carefully planned security structure must be in place. HIPAA mandates biomedical data to be constantly under tight security constraints meeting their proposed standards [78]. Funding agencies frequently require that collaborative biomedical research have a carefully planned security structure [79]. Security structure is not only a necessity due to sensitive nature of the data. It is possible that the mistrust among researchers not wanting to share can be alleviated by tight security constraints [7].

A tight security structure alone, however, is not enough in collaborative research. In collaborative research, certain data are more sensitive than others. It is also possible that some parts of research might only be available to a small number of researchers within the collaboration. To support all levels of constraints, a flexible security structure provides a differential access to a common collaborative system [41,80]. Hence, collaborative security structure at bare minimum involves authentication (minimum login/password) and authorization (variable access to different parts of research data and work area) as well as encryption to ensure communication channels are not weak links as well as audit trails to allow review of who accessed what.

4.2.3. Metadata and data provenance—Before the era of large-scale biomedical collaboration, most research projects occurred within a single discipline, often in a single laboratory. Resulting data were manageable in size and similar in content. The researchers had little difficulty understanding the data since they had similar backgrounds; useful when establishing a common ground or understanding the context of the data. Modern collaborative biomedical research, however, generates large quantities of highly diverse data [71]. Furthermore, interdisciplinary collaborative research involves researchers from various fields. In translational research for example, at one end of the spectrum there might be a bench scientist sequencing genes involved in a particular cancer, while on the other end a clinical researcher may be involved in a clinical trial of a drug for cancer [81]. To facilitate collaborative sharing and analysis, the collaborative research tools must bridge the gap of diversity in data and variability in researchers' field of knowledge. Metadata associated with data sets and data provenance are two ways to bridge that gap [82].

Metadata provides a contextual description of data [83]. Without the context of how research was conducted and data was analyzed, the data are meaningless [75]. For example, say a bench scientist discovers a possible treatment for cancer and gives the result to a clinical researcher. The clinical researcher would find the data impossible to decipher without an explanation associated with the data. Metadata provides such contextual explanation. Data provenance is related to metadata in a sense that it is essentially concerned with the history of data [82], the importance being that it functions like a versioning mechanism. Without a detailed data history, the researchers have no way of determining whether the data is outdated or still valid. Not having data provenance in collaborative research could result in confusion of some researchers using outdated or invalid data while others are using the most updated data. Data provenance enables scientists to keep track of the large data sets generated by collaborative research.

4.2.4. Interoperability and data standards—Interoperability is a concept assumed in most collaboratory research, but rarely explicitly mentioned. Most collaboratory literature implies interoperability by stating how all software involved in collaboration must be able to talk to each other [57] or describes collaboratories as all having a common infrastructure support [40]. Interoperability involves a common infrastructure that seamlessly integrates technology from all levels of research within collaboration. Interoperability is difficult to achieve in collaborative biomedical research due to the heterogeneity of technologies, data formats and content at different levels of research [84]. For example, a genetic sequence of a cancer gene is significantly different from an x-ray image of a tumor. Data standards enable inter-operability by creating a common frame of reference. The researchers from diverse backgrounds can converge on that common frame of reference to exchange of data and share ideas [85–87].

4.3. Review of environmental factors that support collaboratories

Interdisciplinary collaboration does not exist in a vacuum. Collaborations, in particular, are contained within a larger environmental framework such as a university or institution which can have positive or negative effect on the collaborative process. As such, surrounding environmental factors are important to collaborative research. Many studies of collaboratories focus on either the technical or the social issues within the collaboration itself. Very few studies examine the environment surrounding the collaboration (consider [42] as a notable exception). From the literature review of these few [57], three aspects emerged as critical environmental factors that could support collaboration: (1) Administrative and Management structure, (2) Technical support, and (3) Available funding.

4.3.1. Administrative and management structure—The Administrative and Management structure is composed of those involved in collaboration that oversee, coordinate, resolve conflicts, and make decisions. Small collaborations can succeed without a designated administrative body since only a handful of people are involved. The principal investigators involved can essentially function as a managing body. Large interdisciplinary collaborations, however, involve a large number of people and it is often difficult to manage them without a designated administrative body. In these cases, the administrative structure ideally supports the collaborative process by functioning as a central voice for everyone so that no one involved feels isolated [57]. Furthermore, administrative and management structure can also function as an infrastructure to support legal issues [88]. To properly support anything other than very small interdisciplinary collaborations, some body of people has to manage the overall collaborative process [57].

4.3.2. Technical support—Collaborative research often involves complex technology. The researchers not only work with remote communication technologies, but also with tools involved in data sharing and analysis. Since the primary objective of the researchers involved in collaboration is to advance scientific knowledge, they are often less inclined to dedicate time to learning new complex technologies. It is possible that the frustration that comes with such extra work might deter scientists from continuing with the collaboration. Several collaboratories have failed due to such technical complications. For example, the Worm Community System (WCS) was not adopted by some researchers because they had to master relatively complex system installations within unfamiliar computing environments (e.g. the UNIX operating system) [40]. The upper atmospheric collaboratory users were challenged by frequent browser downloads [40]. Hence, interdisciplinary collaboration can benefit by having technical support personnel to alleviate this simple problem of difficulty in technology that could potentially cause collaboration to fail [57].

4.3.3. Available funding—Interdisciplinary collaboration has more financial overhead compared to traditional research to support overall scientific process and technical infrastructure required for coordination of a large group of people, and sharing of data and analysis. To alleviate this difficulty, adequate funding structure should be available as part of the environment fostering collaboration. Olson et al. have found that the collaboration based on funding initiatives are more likely to fail since the lasting collaboration requires something more than a simple financial incentive; however, it would also be difficult for interdisciplinary collaboration to happen without the incentive of funding [57].

5. Future considerations for biomedical collaboration

There is a general consensus that interdisciplinary collaboration is important in solving large-scale, complex biomedical questions. Since the early days of the collaboratory concept, technologies such as communication, and data sharing and analysis tools have been produced to alleviate the problem of remote collaboration, and social issues of collaboration such as willingness to share and trust have been identified and researched. Yet very little is known about how to sustain such collaborations over the longer term. Over the years, some collaboration simply ended when the initial funding ended. Others dissolved due to the overwhelming challenges they faced during the first few years. From our literature review, we identified two factors that might help foster ongoing collaboration over multiple years.

5.1. Appropriate training

The first factor is appropriate training for both the scientific process and technical aspects of collaboration for those involved in the collaborative research. Because technology is important in the daily-to-day functions of an interdisciplinary collaboratory, the difficulties in technology use could negatively effect the resulting collaboration. Additional training in technology for technical support personnel can alleviate the difficulties in technology adoption and use [57].

The knowledge of scientific process itself is also important to collaboration. Due to variability in their background, the researchers involved in collaboration often focus on their local research objectives and overlook the larger goals of the collaborative effort. It is important for all those involved to have an understanding of the overall goal and the high level scientific processes of the entire collaboration. Training in both the technology and scientific process, can help foster collaborative environment through easier understanding of the overall collaboration process and use of technology [57].

5.2. Long-term planning

A second factor important in lasting collaboration is long-term strategic planning. Most interdisciplinary collaborations form out of necessity, either due to a need for expertise or an access to funding resources; hence, these collaborations often form dynamically, as needs arise. Due to their dynamic nature, collaborative efforts seldom plan beyond the first few years, especially not beyond the initial funding period, although the complex problems that these collaborations seek to solve often cannot be expected to be resolved in a few years. Long-term planning of infrastructure and the overall collaborative process is needed from the beginning if the collaborations are to last [57]. Lack of such planning has contributed to the ending of collaboratories before their aims have been achieved [57]. Furthermore, the interdisciplinary collaborative research is a constantly changing process. Supporting collaborative research through continuous evaluation of needs, identifying gaps/barriers, and learning from past successes/failures is imperative for the success of long-term collaboration [61,89,90].

6. Discussion

Although recognized by funding agencies such as NIH as a necessity, interdisciplinary biomedical collaboration is difficult for several reasons. Collaborative biomedical research involves a wide range of expertise from different institutions in disparate locations with little common ground such as a common vocabulary, social norms, or organizational structure. Based on lessons from collaboratories in other disciplines, the implementation of collaborative technology in biomedical research will also need to address this lack of commonality. Furthermore, biomedical collaborations are more likely to fail due to the highly competitive nature of biomedical research [57].

We have introduced many relevant collaboratory concepts in this systematic review. *Communication, common workspace and coordination, and data sharing and management* are general collaboratory concepts essential to biomedical collaboration which all have a biomedical informatics component. *Data integration and analysis, security structure, metadata and data provenance, and interoperability and data standards* are specific collaboratory concepts particularly relevant to biomedical research that are also key areas of research, development and application in biomedical informatics. We identified *administrative and management structure, technical support, and available funding* as three critical environmental factors that could support collaboration. These factors are overarching issues important to all aspects of biomedical research collaborations including the biomedical informatics aspects. We also identified *appropriate training for both the scientific process and technical aspects of collaboration, and long-term planning* as two factors that can continuously support interdisciplinary collaboration.

The collaboratory concepts we reviewed provide ways to guide planning and design of collaborative infrastructure to better support interdisciplinary biomedical collaboration. In particular they have implications for biomedical informatics researchers and developers who are working on ways to support biomedical collaborations. The relevance of these concepts are illustrated in the latest call for proposal of Clinical and Translational Center where it states that biomedical informatics support should consider the following: interoperability, security, workflow, usability and standards [79]. These are only a subset of the concepts we have reviewed. In this paper, we have further extended the list of important concepts biomedical informatics must consider in order to support interdisciplinary collaborative research. It is possible that there needs to be an in-depth study of the work practices of those involved in collaboration as well as technologies involved at each level.

We do not claim that the collaboratory concepts are a panacea for fostering interdisciplinary biomedical collaboration nor for developing useful informatics systems to support it. The goal of this review is not to resolve all difficulties of interdisciplinary biomedical collaboration, but to provide concepts that can help better support interdisciplinary biomedical research collaboration. Furthermore, on the basis of the findings from this review, it is clear that there are benefits for the biomedical informatics community to get involved in collaborative research at every step, particularly early in the process. From the initial formation, biomedical informatics should be part of planning the implementation of collaborative technology to fit the social processes of collaboration. Throughout the collaborative process, informatics support should continuously evaluate and make appropriate adjustments with the ever changing needs and technologies.

In this review, we have focused our efforts on the national science and engineering collaborations; yet interdisciplinary biomedical collaborations that cross international boundaries are beginning to emerge [91]. A whole new set of issues arises when the collaboration crosses national boundaries. These are issues in addition to those we have identified here. Hence the future direction to fostering biomedical collaboration should look towards the interdisciplinary collaboration at the international level.

Summary table

What was already known in topic:

- The technologies that enable remote collaboration were designed and studied by collaboratory researchers in the fields such as Computer Supported Cooperative Work (CSCW) and Information Science (IS).
- The methods to design and evaluate the infrastructure for collaboration were studied in the fields such as CSCW and IS.

What this study added to our knowledge:

- This study brings in some of the prior collaboration researches in other fields such as CSCW and IS into the field of Biomedical Informatics
- This study is a comprehensive review of recurring elements across infrastructure, social and environmental issues that are critical collaborations. Thus far, no existing literature incorporates all the elements into a single review.

Acknowledgments

The authors would like to thank and acknowledge NIH NLM Training Grant (5 T15 LM007442-05) and NIH CTSA Grant (08 URR 025014A) for providing the funding to support parts of this work.

References

1. Hara N, Solomon P, Kim SL, Sonnenwald DH. An Emerging view of scientific collaboration: scientists' perspectives on collaboration and factors that impact collaboration. Journal of the American Society for Information Science and Technology 2003;54(10):952–965.

2. Cummings JN, Kiesler S. Collaborative research across disciplinary and institutional boundaries. Social Studies in Science 2005;35:703–722.
3. Horwitz AR, Watson N, Parsons JT. Breaking barriers through collaboration: the example of the cell migration consortium. Genome Biology 2002;3(11)(pcomment2011)
4. Shrum W, Chompalov I, Genuth J. Trust, conflict and performance in scientific collaborations. Social Studies of Science 2001;31(5):681–730.
5. National Institute of Health Roadmap. http://www.nihroadmap.nih.gov
6. National Meeting on Enhancing the Discipline of Clinical and Translational Sciences. http://www.ncrr.nih.gov/ncrrprog/roadmap/Executive_Summary.pdf
7. Birnholtz, JP.; Bietz, MJ. When Do Researchers Collaborate? Toward a Model of collaboration Propensity in Science and Engineering Research. University of Michigan; Ann Arbor, MI: 2005.
8. Campbell EG, Clarridge BR, Gokhale M, Birenbaum L, Hilgartner S, Holtzman NA, Blumenthal D. Data withholding in academic medicine: characteristics of faculty denied access to research results and biomaterials. Research Policy 2002;29:303–312.
9. Olson, GM.; Teasley, S.; Bietz, MJ.; Cogburn, DL. Collaboratories to support distributed science: the example of international HIV/AIDS research. Proceedings of SAICSIT 2002, ACM International Conference Proceedings Serious; Port Elizabeth, South Africa.
10. O'Day, V.; Annette, A.; Kuchinsky, A.; Bouch, A. Proceedings of the Seventh European Conference on Computer Supported Cooperative Work. Kluwer Academic Publishers; Bonn, Germany: 2001. When worlds collide: molecular biology as interdisciplinary collaboration.
11. Herbsleb, JD.; Mockus, A.; Finholt, T.; Grinter, RE. Proceedings of Computer Supported Cooperative Work 2000. ACM; New York, NY, US: 2000. Distance, dependencies, and delays in a global collaboration.
12. Kraut, R.; Egido, C. Patterns of contact and communication in scientific research collaboration. Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work; Portland, OR, USA. 1990.
13. Commission on Systematic Interoperability: Ending the Document Game. http://www.endingthedocumentgame
14. National Health Information Infrastructure. http://www.aspe.hhs.gov/sp/NHII
15. Addis, M.; Ferris, J.; Greenwood, M.; Li, P.; Marvin, D.; Oinn, T.; Wipat, A. Experiences with e-Science work-flow specification and enactment in bioinformatics. Proceedings of UK e-Science All Hands Meeting; Nottingham, UK. 2003.
16. Arnstein, LF.; Borriello, G.; Consolvo, S.; Hung, C.; Su, J. Labscape: a smart environment for the cell biology laboratory; IEEE Pervasive Computing Magazine; 2002. p. 13-21.
17. Arnstein, LF.; Grimm, R.; Hung, C.; Hee, J.; LaMarca, A.; Sigurdsson, SB.; Su, J.; Borriello, G. Systems support for ubiquitous computing: a case study of two implementations of Labscape. Proceedings of the First International Conference on Pervasive Computing; Springer-Verlag, Germany. 2002.
18. Ash JS, Gorman PN, Lavelle M, Payne TH, Massaro TA, Frantz GL, Lyman JA. A cross-site qualitative study of physician order entry. Journal of American Medical Informatics Association 2003;10(2):188–200.
19. Bafoutsou G, Mentzas G. Review and functional classification of collaborative systems. International Journal of Information Management 2002;22(4):281–303.
20. Baker, KS. Ecological design: an interdisciplinary, interactive participation process in an information environment. Proceedings of the workshop on Requirements Capture for Collaboration in e-Science; Edinburgh, UK. 2004.
21. Baker, KS.; Jackson, SJ.; Wanetick, JR. Strategies supporting heterogeneous data and interdisciplinary collaboration: towards an ocean informatics environment. Proceedings of the 38th Hawaii International Conference on System Science, vol. HICSS38; Big Island, HI, US. 2005.
22. Bly, S.; Keith, KM.; Henline, PA. The work of scientists and the building of collaboratories. Proceedings of the Group 97 International Conference on Supporting Group work, 1997, General Atomics Report GA-A22619; Phoenix, AZ, US. (Preprint)
23. Casper TA, Meyer WM, Moller JM. Collaboratory operations in magnetic fusion scientific research. Interactions 1998;3(56):56–64.

24. Chin G, Myers J, Hoyt D. Social networks in the virtual science laboratory. Communications of the ACM 2002;45(8):87–92.

25. Finholt TA, Olson GM. From Laboratories to Collaboratories: A New Organizationalform for Scientific Collaboration 1997;8(1):28–36.

26. Finholt, TA. Collaboratories: science over the internet. In: Teich, Albert H.; Nelson, Stephen D.; Lita, Stephen J., editors. AAAS Science and Technology Policy Yearbook. American Association for the Advancement of Science; Washington, DC: 2002. p. 339-344.

27. Fisher, D.; Dourish, P. Social and temporalstructures in everyday collaboration. Proceedings of the SIGCHI Conference on Human factors in computing Systems; Vienna, Austria. 2004.

28. Foster, I.; Vockler, J.; Wilde, M.; Zhao, Y. The virtual data grid: a new model and architecture for data-intensive collaboration. Proceedings of the First CIDR—Biennial Conference on Innovative DataSystems Research; Asilomar, CA, US. 2003.

29. Halkola, E. Scientific Collaboration and Information Infrastructures—Information Practices in LIDET Experiment. 2004.

30. Jeffrey P. Smoothing the waters: observations on the process of cross-disciplinary research collaboration. Social Studies of Science 2003;33:539–562.

31. Katz JS, Martin BR. What is research collaboration. Research Policy 1997;26:1–18.

32. Melin G. Pragmatism and self-organization: research collaboration on the individual level. Research Policy 2000;29:31–40.

33. Myers, JD.; Allison, TC.; Bittner, S.; Didier, B.; Frenklach, M.; Green, WH., et al. A collaborative informatics infrastructure for multi-scale science. The Proceedings of the Challenges of Large Applications in Distributed Environments (CLADE) Workshop; Honolulu, HI, USA, IEEE Computer Society. 2004.

34. Olson, JS.; Teasley, S. Groupware in the wild: lessons learned from a year of virtual collocation. In Proceedings of the 1996 ACM Conference on Computer-Supported Cooperative Work (CSCW'96); ACM Press. 1996.

35. Schleyer T. Collaboratories: leveraging information technology for cooperative research. Journal of Dental Research 2001;80(6):1508–1512. [PubMed: 11499503]

36. Sonnenwald, DH.; Bergquist, RE.; Maglaughlin, KL.; Soo, EK.; Whitton, MC. Designing to support scientific research across distances: the nanomanipulator environment. In: Churchill, DSAMEE., editor. Collaborative Virtual Environments. Springer-Verlang; London: 2001.

37. Teasley S, Wolinsky S. Scientific collaborations at a distance. Science 2001;292(5525):2254–2255. [PubMed: 11423638]

38. Wagner CS. Six case studies of international collaboration in science. Scientometrics 2005;62(1):3–26.

39. Sonnenwald DH, Whitton MC, Maglaughlin KL. Evaluating a scientific collaboratory: results of a controlled experiment. ACM Transactions on Computer-Human Interaction 2003;10(2):150–176.

40. Finholt TA. Collaboratories as a new form of scientific organization. Economics of Innovation and New Technology 2003;12(1):5–25.

41. Kouzes RT, Myers JD, Wulf WA. Collaboratories: doing science on the internet. IEEE Computing Practices 1996;29:40–46.

42. Star, SL.; Ruhleder, K. Steps towards an ecology of infrastructure: complexproblems in design and access for large-scale collaborative systems. Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work; ACM Press, Chapel Hill, NC. 1994.

43. Worm Community System. http://www.canis.uiuc.edu/projects/wcs/index.html

44. Olson GM, Atkins DE, Clauer R, Finholt TA, Jahanian F, Killeen TL, Prakash A, Weymouth T. The upper atmospheric research collaboratory. Interactions 1998;5(3):48–55.

45. Environmental Molecular Sciences Laboratory. http://collaboratory.emsl.pnl.gov

46. Schur A, Keating KA, Payne DA, Valdez T, Yates KR, Myers JD. Collaborative suites for experiment-oriented scientific research. ACM Interactions 1998;3(5):40–47.

47. Bhandarkar, M.; Budescu, G.; Humphrey, W.; Izaguirre, JA.; Izrailev, S.; Kalé, LV.; Kosztin, D.; Molnar, F.; Phillips, JC.; Schulten, K. BioCoRE: a collaboratory for structural biology. Proceedings

of the SCS International Conference on Web-based Modeling and Simulation; San Francisco, CA, US. 1999.

48. Biological Collaborative Environment. http://www.ks.uiuc.edu/Research/biocore, BioCoRE
49. Biological Sciences Collaboratory. http://www.pnl.gov/biology, BSC
50. Chin, G.; Lansing, CS. Capturing and supporting contexts for scientific data sharing via the biological sciences collaboratory. Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work; Chicago, IL, US. 2004.
51. Bødker, S.; Grønbæk, K.; Kyng, M. Participatory Design: Principles and Practices. Lawrence Erlbaum Associates; Hillsdale, NJ: 1993. Cooperative Design: Techniques and Experiences from the Scandinavian Scene; p. 157-175.
52. Muller MJ, Wildman DM, White EA. Taxonomy of PD practices: a brief practitioner's guide. Communications of the ACM 1993;36(4):26–27.
53. Sumner, T.; Stolze, M. Evolution, Not Revolution: Participatory Design in the Toolbelt Era. In: Mathiassen, MKaL, editor. Computers and Design in Context. MIT Press; Cambridge, MA: 1997. p. 1-26.
54. Weng C, McDonald DW, Sparks D, McCoy J, Gennari JH. Participatory design of a collaborative clinical trial protocol writing system. International Journal of Medical Informatics 2006;76(S1):245–251.
55. Beyer, H.; Holtzblatt, K. Contextual Design: Defining Customer-Centered Systems. Morgan Kaufmann Publishers; 1998.
56. Winograd T. Categories, disciplines, and social coordination. CSCW 1994;2:191–197.
57. Olson, JS.; Olson, GM.; Hofer, EC. What makes for success in science and engineering collaboratories?. Workshop on Advanced Collaborative Environments; Redmond, WA, US. 2005.
58. Google Scholar. http://scholar.google.com
59. PubMed Central. http://www.pubmedcentral.nih.gov
60. Hollan, JD.; Stornetta, S. Beyond being there. Proceedings of'ACM SIGCHI'92; Monterey, CA, US. 1992.
61. LaCoursier, S.; Sarkar, M. Medinfo 2004. 2004. Communication and information needs and barriers: an international collaboration model.
62. Shortliffe EH, Patel VL, Cimino JJ, Barnett GO, Greenes RA. A study of collaboration among medical informatics research laboratories Artificial intelligence in medicine. 1998;12(2):97–123.
63. Dourish, P.; Bellotti, V. Awareness and coordination in shared workspaces. Proceedings of CSCW'92, ACM Press; Toronto, Canada. 1992.
64. Gennari JH, Weng C, Benedetti J, McDonald DW. Asynchronous communication among clinical researchers: a study for systems design. International Journal of Medical Informatics 2005;72(10): 797–807. [PubMed: 16023408]
65. Bannon, L.; Bodker, S. Constructing common information space. Proceedings of the European Conference on Computer-Supported Cooperative Work ECSCW'97; Kluwer Academic Publishers, Lancaster, UK. 1997.
66. Ackerman MS, McDonald DW. Collaborative support for informal information in collective memory systems. Information Systems Frontiers 2000;2(3–4):333–347.
67. Appelt, W. WWW based collaboration with the BSCW system. Proceedings of SOFSEM'99, Springer Lecture Notes in Computer Science; Milovy, Czech Republic. 1999. p. 1725
68. Hoffelner, W. Report from ETH Zürich. Applicability of of groupware for communication in different project environments—a case study.
69. Wagner CS. Wiki: a technology for conversational knowledge management and group collaboration. Communications of the Association for Information Systems 2004;13:265–289.
70. Basic Support for Cooperative Work. http://www.bscw.fit.fraunhofer.de/, BSCW
71. Ball CA, Sherlock G, Brazma A. Funding high-throughput data sharing. Nature Biotechnology 2004;22(9):1179–1183.
72. Birnholtz, JP.; Bietz, MJ. Data at work: supporting sharing in science and engineering. Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work (GROUP'03); Sanibel Island, FL, US. 2003.

73. NIH data sharing policy. http://grants.nih.gov/grants/policy/data_sharing
74. NSF data sharing policy. http://www.nsf.gov/pubs/2001/gc101/gc101rev1.pdf
75. Allison DB, Cui X, Page GP, Sabripou M. Microarray data analysis: from disarray to consolidation and consensus. Nature Reviews Genetics 2006;2000:55–65.
76. Altman R. Share and share alike: a proposed set of guidelines for both data and software. Biomedical Computer Review (Summer) 2006:1.
77. Louie B, Mork P, Martin-Sanchez F, Halevy A, Tarczy-Hornoch P. Data integration and genomic medicine. Journal of Biomedical Informatics 2007;40(1):5–16. [PubMed: 16574494]
78. Health Insurance Portability and Accountability Act. http://www.cms.hhs.gov/HIPAAGenInfo
79. Clinical and Translational Science Awards RFA. http://www.ncrr.nih.gov/clinicaldiscipline.asp
80. Blobel B. Comparing approaches for advanced e-health security infrastructures. International Journal of Medical Informatics 2007;76(5–6):454–459. [PubMed: 17074532]
81. Mankoff SP, Brander C, Ferrone S, Marincola FM. Lost in translation: obstacles to translational medicine. Journal of Translational Medicine 2004;2(1):14. [PubMed: 15149545]
82. Simmhan Y, Plale B, Gannon D. A survey of data provenance in e-science. SIGMOD Record 2005;34 (3):31–36.
83. Buetow K. Cyberinfrastructure:, Empowering a "third way" in biomedical research. Science 2005;308 (5723):821–824. [PubMed: 15879210]
84. Maojo V, Kulikowski CA. Bioinformatics, Medical informatics: collaborations on the road to genomic medicine? Journal of the American Medical Informatics Association 2003;10(6):515–522. [PubMed: 12925552]
85. Bruun-Rasmussen M, Bernstein K, Chronaki C. Collaboration—a new IT-service in the next generation of regional health care networks. International Journal of Medical Informatics 2003;70 (2–3):205–214. [PubMed: 12909171]
86. Neches R, Fikes R, Finin T, Gruber T, Patil R, Senator T, et al. Enabling technology for knowledge sharing. AI Magazine 1991;12(3):36–56.
87. Häkkinen H, Korpela M. A participatory assessment of IS integration needs in maternity clinics using activity theory. International Journal of Medical Informatics 2007;76(11–12):843–849. [PubMed: 17174147]
88. David, P.; Spence, M. The Institutional Infrastructure of e-Science: The Scope of the Issues. Towards institutional infrastructures for e-Science: the scope of the challenge, 2003, Final Report of the Oxford Internet Institute project.
89. Kaplan B, Shaw NT. Future directions in evaluation research: people, organizational, and social issues. Methods of Information in Medicine 2004;43(3):215–231. [PubMed: 15227551]
90. Sonnenwald, DH. Expectations for a scientific collaboratory: a case study. Proceedings of the 2003 International ACM SIGGROUP, 2003, ACM Press; Sanibel Island, FL, US.
91. Luukkonen T, Persson O, Sivertsen G. Understanding patterns of international scientific collaboration. Science Technology & Human Values 1992;17(1):101–126.

Table 1

A summary of collaboratory concepts and definitions

		Collaboratory features	Definition
General collaboratory concepts across many domains	Communication	Communication-asynchronous	A method of communication that enables researchers to communicate non-simultaneously (e.g. e-mail, discussion list, newsgroup)
		Communication-synchronous	A method of communication that enables researchers to communicate simultaneously (i.e. phone, videoconferencing, whiteboard, chat)
		Awareness	Awareness is an understanding of the activities of others, which provides a context for your own activity
	Common workspace, coordination	Common workspace	A physical or a virtual space where researchers can interact and work together(i.e. web portals, wiki)
		Coordination tools	Tools that help manage scheduling and coordination of various tasks involved in collaboration (i.e. scheduler, calendar)
	Data sharing and management	Data management	Technologies that effectively help handle (i.e. retrieve, search, access) data created during research
		Data sharing	Any data produced during research being shared as well as technologies that are involved in sharing
		Archive/repository	A place where researchers can put their data to store, retrieve, and access
		Trust	A trust is a belief in the integrity and ability of others involved in the collaboration, which facilitate collaborators to work together
		Willingness to share	A willingness is an inclination and readiness to be part of collaboration and to share in the process of a common research goal
Specific collaboratory concepts of particular biomedical relevance	Data integration, analysis	Data integration	A mechanism to integrate data that are in multiple formats and located in different places
		Data analysis tools	Tools that help analyze data in any way (i.e. visualization, display, statistics)
	Security	Security, variable access	A measure to secure access to a system or data only to those authorized to use it
	Metadata, data provenance	Metadata (annotations)	Contextual information, descriptions about a dataset
		Data tracking (data provenance)	Tools to help keep track of history of a dataset
	Interoperability, data standards	Common vocabulary/standards	A common set of defined words or formats for a data set
		Interoperability	All the technologies in a collaboration integrated to work together and interconnected

		Collaboratory features	**Definition**
		Collaboratory features	Definition
Environmental factors that support collaboration	Administration, management	Management structure	A body of people responsible for managing the overall structure of collaboration
	Technical support	Technology support	Having personnel that help with technology(i.e. set up new hardware, software support)
	Funding	Funding	Funding for collaboration, incentive
Continuous support of collaboration	Training	Education/training	Any tools and activities related to educating, mentoring, and training
	Iterative evaluation	Iterative evaluation	Continuous evaluation of needs and barriers throughout the collaborative process (both technological and social)
		User-centered design	Designing tools with users in mind (i.e. participatory design, ethnographic study of workflow)
		Understanding workflow	Understanding the overall structure of work of those involved in collaboration