Executive Summary

Our committee makes three main recommendations:

1) CSAIL should continue its tradition of providing a Reading Room to house books, provide quiet study space for those without individual offices, and create a comfortable venue for browsing and reading journals, conference proceedings, magazines and newspapers. As technical reports, journals and bibliographic databases increasingly become available on-line, we need not retain physical copies of titles that are accessible through MIT’s or others’ on-line sites. The Reading Room should also serve as a venue for pick-up and drop-off of material borrowed from MIT Libraries and its services. We recommend that Reading Room space be reconfigured to provide more place to read by moving some older reports and journals to storage, and that new, more comfortable seating should replace the disintegrating furniture. We recommend that the Reading Room be staffed and open during some evening hours, and perhaps open 24 hours/day with cardkey access.

2) We should create a new CSAIL Archivist position that subsumes the previous Librarian position. We envision the Archivist as a peer to the PI’s, attending PI meetings and the annual retreat, and becoming broadly familiar with the work of the Lab. The Archivist would be charged with tracking the intellectual input and output of Lab members over time, and with better integration of the Reading Room collections, member Publications, and exhibition of members’ creations. The Archivist would take an active role in promoting use of the Reading Room and disseminating information about new material. The Archivist would also be a shaper of and conduit for the Lab’s recurring interactions with MIT Libraries, giving us a more consistent voice.

3) The Lab should implement a common collection of electronic versions of conference proceedings, paper reprints and other works to minimize duplicated effort by different groups and encourage sharing. We should also exploit and develop technology that supports scanning, indexing and retrieval of such resources with minimal effort, create or adopt means to “fall back” from our local repository to material available through MIT Libraries or other licensed sources, and use our collection as the basis for “social networking” applications that indicate to students what faculty and others consider worth reading and discussing.

Process

Prof. Brooks had appointed our committee in late summer 2006 to report a set of options about how CSAIL should address the following issues:

- Should we have a physical reading room of the sort that has been traditionally maintained by the Lab and its predecessors for decades? If so, where should it be located, what services should it provide, and what steps should the Lab take to make sure that the usage of the reading room is commensurate with our investment in its maintenance?
What should be the relationship between a CSAIL reading room and the services provided by the MIT Libraries? We also considered a generalization of this question to include also the MIT Archives and CSAIL publications.

Based on the dramatic changes to information storage and retrieval technologies and corresponding changes in people’s use of information resources, what current practices need no longer be pursued by the reading room, and what new capabilities could we provide to improve the intellectual environment of our faculty, staff and students?

The members of the committee are:

- Peter Szolovits (co-chair)
- Jack Costanza (co-chair)
- Michael Collins
- Tom Knight
- Ronitt Rubinfeld
- Karen Sollins
- Seth Teller
- Aaron Adler (graduate student)

To address these issues, the committee has held frequent meetings since late September, where we have interviewed relevant people from within and outside the Lab. Subsets of us have also spoken with others who were unable to attend our meetings. In these ways, we have spoken with

- Rod Brooks, who charged the committee
- Ann Wolpert, head of MIT Libraries
- Amy Stout, MIT’s Computer Science librarian
- MacKenzie Smith, Associate Director for Technology of the MIT Libraries
- Danny Weitzner, who provided some general legal directions on IP issues
- Paula Mickevich, who staffs the current reading room
- John Guttag, who is responsible for space allocations in the Lab’s CSAIL territory
- Hal Abelson, who has been leading discussions around the Institute about data sharing
- David Karger, whose research interests include personally organized but shared electronic information resources
- Frans Kaashoek and Jeremy Stribling, who have built and supported bibliographic search software

We made a presentation at a Thursday Lab lunch to solicit the views of our colleagues and at a later one to outline some of the ideas presented here and to invite our colleagues to participate in an on-line survey. We also spoke extensively, though informally, with members of our research groups and groups of graduate students in order to understand the ways in which they currently use the reading room and what resources they rely on. We also received feedback from students at GSL and GSB.

Finally, we conducted a survey of members of the Lab, whose results have validated some of our ideas and toned down others. Of a total of 900 individuals on the csail-all email list who received invitations to take the survey, 192 did so, for a respectable response rate of 21%. Of the 192 respondents, 29 were faculty, 27 researchers, 22 staff, 102 doctoral
students, 7 SM/MEng students, 2 undergrads and 3 visitors. All but a final question about funding levels asked respondents to rate ideas on a modified, asymmetric Likert scale, with possible responses “1—Bad Idea”, “2—Neutral”, “3—Desirable”, “4—Highly Desirable” and “5—Essential”. We asked them to restrict their “Essential” and “Highly Desirable” votes to no more than eight (1/4 of the questions) so we could use ratings to estimate relative preferences. In aggregating these data, any number above 2.0 is favorable. Of the suggestions we explored in the survey, only one received an average score below 2.0; 13 were rated between 2.0 and 3.0, 16 between 3.0 and 4.0, and two above 4.0. In general, ratings by faculty and researchers ran about 0.2 less favorable than ratings by students and non-research staff, and we comment in places on some apparent systematic differences of view between these two subgroups.

A Physical Reading Room

We surveyed many of our colleagues at other computer science departments and found that few had dedicated reading rooms, and most that did have such rooms had recently either discontinued them or turned them into lounges, game rooms, coffee bars, or other social spaces. Nevertheless, both in our discussions with students and in the survey, we found a very strong desire to maintain an actual reading room in CSAIL. The average survey rating for this question was slightly above 4.0. Faculty and researchers gave it only 3.5, but students and others rated it at 4.3, tied for their highest rating. The two main themes that we encountered again and again were the availability of many useful books that often did not exist within research groups and the need for a quiet place to retreat when shared offices or cubicles became distracting. People also found the natural light and views attractive, and highly favored comfortable seating. Indeed, we recommend that additional comfortable reading chairs should be purchased to replace those now falling apart and to expand the number of seats.

Even among faculty and researchers, the largest number of respondents (17) considered a physical reading room essential, 14 and 11 voted it desirable or highly desirable, and only 6 thought it a bad idea. Among others, 75/130 (58%) found it essential, and only 2 a bad idea.

The committee and most respondents also favor making the reading room available to CSAIL members during extended evening hours or around the clock, though we also see concerns about the possible volatility of the collection, and would hesitate to propose spending for 24 hour security. We believe that viable options to pursue here include hiring students to keep the room open until, say, 8pm on weekdays, or experimenting with greater access via prox cards, perhaps accompanied by recorded video surveillance, to determine if pilferage turns out to be a problem. We note that paying a student $15/hour, 3 hours/workday, for 220 workdays/year comes to about $10,000 (without overhead). That amount is approximately the cost of replacing 100 books, in a purely utilitarian calculation.

We also discussed ways in which the collection and circulation processes of our reading room might be better integrated with those of MIT Libraries. Our recommendation is that tighter integration should not be pursued at this time. For example, there are potentially serious policy issues surrounding the financing of the reading room that might be brought to a head by integrating it with MIT Libraries’ circulation system. In particular, we were warned that the Institute would not take on funding of the Reading Room, that its current funding as part of the CSAIL allocation is justified only by its exclusive service to CSAIL, and that the MIT circulation system would not support such limitations. In addition, we got the sense that the systems used by MIT Libraries were too “heavy weight” for our needs. We therefore encourage CSAIL to modernize and augment our cataloging, search and circulation systems along the lines that TiG has begun to

---

1 The survey asked participants to select among these mutually exclusive categories. Some also identified themselves. The numbers reported here include two researchers who categorized themselves as staff but, from their identities, are researchers, and one faculty member who did not select a category. Because some misunderstood the staff category, it may contain a few more researchers who characterized themselves as staff, but because they did not identify themselves, we cannot tell.
investigate, based on open-source software more appropriate to small operations such as ours. We do believe that we should arrange for our reading room to serve as a pick-up and drop-off point for materials available through MIT Libraries. (Because there is already a drop-off bin on the Student Street in Stata, pick-up is the more important function.)

Our committee began with considerable enthusiasm for finding the ideal space in which to locate a reading room in the Stata Center, but we were discouraged from this task by two factors. First, Prof. Guttag, who heads the Lab’s space committee, made clear to us that such decisions were in the purview of his committee, not ours, and that many considerations outside our discussions constrained space decisions. Second, after touring potential locations for a reading room in the building, we did not find a compellingly attractive solution. Our favored recommendation was to rebuild space along both sides of the fourth floor corridor connecting the R&D to Building 36 to create a complex that could at once:

• house a collection of books, etc.;
• provide an archival space for collection and display of an interesting array of artifacts from the Lab and thus serve as a kind of mini-museum, public exhibition space and visitors’ center;
• be proximate to lounge space, game tables, and refreshments, as well as the outside patio during good weather; and
• attract attention from the large flow of traffic through the area from other parts of the Institute to faculty dining, the pub and CSAIL headquarters.

Our initial impression was that such a confluence of potential uses of the space would increase its utilization, and if subspaces could be sound-isolated but visually connected, this could provide a vibrant center of activity that would help to overcome the current reading room’s relative isolation and thus draw more people to it.

However, both from discussions and from our survey, we saw significant resistance to this proposal. For example, our question about locating a reading room in a high-traffic area received a rating of 2.2, third lowest of 31 questions. Interestingly, in a reversal of the usual pattern, faculty and researchers rated this suggestion much (0.7 points) higher than others did, among whom 54 (41%) rated it a bad idea. We also found only tepid support for maintaining a collection of artifacts, a display space, or integration with social uses such as lounges, cable TV and coffee. It may be that the availability of many other spaces in the building suitable for such use and the very strongly expressed desire for the reading room to provide quiet space, which students view as antithetical to “high traffic,” account for this rejection. We also sensed some innate conservatism, which may account for the favorable views of the current reading room’s location, views, natural light and quiet.

Physical Holdings

We found that members of our community consider the approximately 3,000 books held by the reading room to be its most beneficial resource. In addition, people often browse newspapers, general circulation magazines and some of the most popular journals in the reading room, and a few have suggested adding subscriptions to additional wide-scope journals such as Nature. We also found that there is only very limited use of technical reports and the somewhat fragmented collections of past journal issues on the shelves. In general, it is our recommendation that our reading room should not retain copies of old material that is available on-line. However, because many journals still have not digitized their older back issues or because MIT may not subscribe to such collections, the most worthwhile such material to retain is just that collection of journals that are not otherwise conveniently available. Nevertheless, it may be appropriate to house such collections in more densely organized spaces that are commensurate with the frequency with which people need to access them. Moving these collections from the reading room to storage space should then free up attractive space for other uses, such as additional seating.
The committee also explored alternative means of holding a large and usable collection, and developed considerable enthusiasm for a distributed scheme that integrates the existing and ongoing holdings of the faculty, researchers and students. We considered several graded steps toward this goal:

1. Create a joint catalog of all the books held by members of the Lab, with each individual’s consent. This would permit others in the Lab to find material they need and to make private arrangements to borrow it, with mutual consent.

2. Create a more formal circulation system that supports easy check-out of books from others’ shelves and tracking and recall of borrowed material.

3. Expand the reading room’s current policy of purchasing books needed by individuals to do so on a larger scale and to have such newly acquired books shelved in the requestor’s collection, if he or she consents to others borrowing it.

4. Fully distribute the reading room’s existing collection of books (and journals) to individuals’ labs and offices.

Both the survey and other responses indicated limited enthusiasm for these approaches. These suggestions received scores of 2.7, 2.9, 2.4, and 1.8, respectively. The suggestions to distribute either new or existing books received among the five lowest ratings of any suggestion, and in both cases “Bad Idea” led or tied with all other votes. Numerous people commented on the possible defects of such a distributed system, and Russ Cox wrote of his experience working last summer at Google, where they tried and abandoned a similar system during his stay. The creation of a joint catalog and check-out system were not as unpopular, and both received more desirable or higher votes than neutral or negative. Nevertheless, there were 30 and 29 “Bad Idea” votes for these suggestions, compared with 36+12 and 34+20 highly desirable and essential votes. Because participation would be purely voluntary, we recommend proceeding with the first two of these recommendations (but not the ones to distributed the reading room’s collection); those who object can simply choose not to participate. We also note that some “private” book collections are in open lab spaces whereas others are in the typically locked offices of individuals. A permissive check out system is much easier to implement for the former.

We also discussed the possibility of bar-coding both the Reading Room’s collection of books and (consistently with the first two recommendations above), also offering that service to individuals in the Lab. We envision this task being concurrent with cataloging. Having these bar codes should simplify and speed up check-out procedures. Visually prominent bar codes will also make it easy to spot Reading Room or shared holdings around the Lab. We also recommend that a small set of valuable but hard-to-find items, to be identified by individual Lab members, be scanned and hosted locally, with access restricted to Lab members.

**Electronic Holdings**

The advent and spread of electronic documents underlies much of the discussion of whether we need a reading room and what information services it or other facilities should provide. For example, thirty years ago the only rapid way we could read technical reports from sister institutions was to find paper copies in our reading room, provided as part of cooperative agreements we had with many other institutions to share such reports. Today, virtually all reports are on-line, and it is much easier to find them on a web site, through Google Scholar or other search engines than to rely on any physical collection. Nevertheless, access to electronic documents is still awkward and much more difficult than we would wish. Although technical reports are freely available, conference proceedings and papers are usually provided only on CD-ROM or perhaps on the web site of the sponsoring organization, which may make them accessible only to
members. Journal articles are particularly problematic because of changing policies of publishers, institutions, funding agencies, etc. As a result, finding a journal article if one runs across a reference to it may be as simple as a couple of links clicked in a Web browser to reach the author’s (legal) posting of a pre-publication version or it may be as difficult as having to traverse the arcane ILLiad system of MIT Libraries to locate and request a scanned version of a paper from a journal to which MIT does not have on-line access. In the latter case, access can be delayed by several days or more.

We propose a significant effort to develop a comprehensive system for giving Lab members access to the published literature in as simple and uniform a manner as possible. The initial goal would be to create a local Lab-wide repository of papers contributed by any members of the Lab. These could include conference papers imported from proceedings CD-ROMs, electronic copies of papers acquired from other sources such as Web searches or downloads from the Library or other sites, and locally-made scans of papers for which no electronic source is available. The Reading Room “Archivist” (see below) would provide an individual point of contact to members wishing to either deposit materials to the shared archive, or wishing to retrieve such materials. Russ Tedrake pointed out the existence of such a local repository maintained by our robotics researchers: https://groups.csail.mit.edu/robotics-center/elib/. Other likely useful candidate technologies to support this repository include the Citeseer mirror (http://citeseer.csail.mit.edu/) and OverCite (http://www.overcite.org:13070/) projects from Frans Kaashoek’s group and various efforts within the Simile project (http://simile.mit.edu/) that is jointly conducted by David Karger’s Haystack project and the MIT Libraries.

The committee noted that many valuable information resources are available through the MIT Libraries, but in so many incompatible ways that finding them is a daunting task. For example, our impression is that Barton (the catalog), Vera (the on-line journal subscription system) and the growing set of open-source journals are completely unrelated. Therefore, one may need to consult all three, using different approaches and interfaces, in search of a particular item. Software used for cataloging contains very valuable information, but access to that is restricted both by technical limitations and restrictive agreements. Discussion with MacKenzie Smith suggests that none of the MIT software is easily amenable to interoperation. MIT Libraries have plans to improve these search mechanisms, and if they succeed, we should exploit them in our solutions. Other existing but un-integrated routes to valuable information come through professional organization sites that provide access through memberships and publishers’ sites that are open to those with individual subscriptions.

The state of the art system that we envision takes a user-centric, not supplier-centric viewpoint, and retrieves references, citations and documents from all available sources. This goal goes back at least to Vannevar Bush’s Memex paper, and has motivated subsequent work in each decade, including Frank Reintjes’s Intrex project in our department and the current efforts mentioned above. Although advances in technology make this goal seem more reachable, we are not under the illusion that it is easy. Therefore we propose a graded effort, first to provide a Lab-wide system such as those already in existence for other groups, and gradually incorporating more advanced methods that can successfully reach out to all available information sources. This will require careful management of development and expectations, and a realistic assessment of what can be achieved with limited resources. The proposed effort can, however, answer the question of how CSAIL could contribute to creation of a model state of the art reading room, which was one of Rod Brooks’ motivating questions for this committee.

We also discussed the potential legal morass that creation of such a shared information resource could engage us in. Unfortunately, many of the specific questions of copyright law and intellectual property potentially raised here are not clearly answered, according to Danny Weitzner. On the theory that it’s better to say “sorry” than to ask and be refused permission, we believe that we should proceed, prohibit access to our system from those outside CSAIL, and only modify our plans if potential legal threats actually materialize.
Seth Teller’s Publications Committee recommended over two years ago that CSAIL should organize a repository of our own products, to be hosted in DSPACE but with a CSAIL front end that would provide to outsiders a well-organized view of the material we have produced, and to insiders an archival repository of our own proposals, reports, documentation, programs, etc. Because of technical and policy restrictions in DSPACE, this plan has so far succeeded mainly for archiving and dissemination of our own reports and theses. We note that, except for management of access by outsiders, many of the desiderata of the effort we propose here overlap with those of the publications effort. Annika Pfluger, in her comments responding to the survey, suggests that “CSAIL Publications and the Reading Room should be more closely linked together since our basic functions are similar.” She has also considered implementing some of the kinds of “new millennium” facilities we recommend, and wants to participate in the proposed project.

In light of the convergence between the functions of library and publications, and our perception of the need to better archive our output, we propose creation of a role, perhaps to be called Archivist, that integrates these responsibilities. Creating a unity of administrative structure may help achieve the better integration of these functions that we believe is possible.

Resources

We were told by Karen Shirer that the current budget of the reading room is $162,000 per year. We believe that the current level of spending is significantly below this amount, because that budget is probably based on the former two staff members working in the reading room and a more active acquisition process than has operated recently. We did not feel empowered to suggest to the Lab leadership how much you should spend on this activity, but the amount to be spent clearly has a decisive influence on which of our other recommendations can be accomplished. We did want to reflect the sense of the Lab about this question, however.

In our survey, with some trepidation, we asked whether people would prefer to spend an amount that is a “1—significant decrease”, “2—decrease”, “3—present amount”, “4—increase”, or “5—significant increase” compared to the estimated baseline of $160,000 that we mentioned. The reason for our trepidation was that such questions historically sometimes elicit “tax revolt” responses. As Rod has pointed out on occasions, “it’s your money, not the Lab’s.” Therefore, it came as somewhat of a surprise that average response was 3.2±0.8, showing a slight preference to spend more than currently. Among faculty and researchers, who actually pay the bills, the result was 2.9±0.9, reflecting a distribution wherein 21 of 53 (40%) responses were for the present amount, 15 (28%) were in favor of increased spending, 12 (23%) favored decreased spending, and 5 (9%) favored a significant decrease. Not surprisingly, favoring a strong decrease correlated strongly with not wanting a reading room at all.

Overall, therefore, there seems to be a willingness in the Lab, both among faculty and researchers and among students and non-research staff to continue to pay for the reading room and its services at approximately the currently budgeted levels.

Implementation and Oversight

We recommend that a small number of PIs who are dedicated to the Reading Room be charged with overseeing the implementation of selected recommendations, in collaboration with TiG and with the Lab’s Executive and Space Committees. We recommend that the Lab incorporate an explicit Reading Room oversight function into one of its existing committees, for example the Executive Committee, in order to monitor Reading Room usage and issues, and Lab members’ satisfaction with the Reading Room, regularly over time (we suggest quarterly review).