New Initiatives in Chemical Education

An On-Line Computer Conference, Summer 1996

Paper 09

USE OF THE INTERNET IN TEACHING CHEMICAL INFORMATION COURSES

Final version: 30 May 1996

GARY WIGGINS
INDIANA UNIVERSITY
CHEMISTRY LIBRARY
BLOOMINGTON, INDIANA 47405 USA
wiggins@indiana.edu

Abstract

Beginning in the Fall semester 1994, a series of chemical information courses was taught at Indiana University via the Internet. The initial attempt utilized LISTSERV as the delivery mechanism, followed by gopher in the Spring 1995 semester, and the Web in the Fall of 1995. Experiences with each of the Internet tools are presented in this paper, and an assessment of the success of the endeavor is made. The advantages of such an approach are summarized, and the rich sources available on the Internet for teaching chemical information are surveyed.

Section 0

0. Table of Contents

Section I

I. Use of the LISTSERV Program for C400: Fall 1994

In September 1994, the CICOURSE (Chemical Information Courses on the Internet) listserve was established at Indiana University. The original purpose of CICOURSE was to provide a forum in which

anyone could teach and learn about chemical information sources. CICOURSE was open to both students and instructors. It was hoped that others might offer to share their expertise in the chemical information area through short courses, workshops, etc. In fact, the C400 Chemical Information Sources and Services course at Indiana University was the only course ever provided via CICOURSE.

Although the number of subscribers to CICOURSE has remained relatively stable (about 300-350), the very low level of traffic since the initial offering led in 1996 to a change in mission. Now CICOURSE exists simply to announce the availability of new or revised chemical information materials that have been made available on servers at Indiana University. For those who wish to see the list of files on CICOURSE, send the one-line message GET CICOURSE FILELIST to LISTSERV@IUBVM.UCS.INDIANA.EDU. No attempt to maintain the files on LISTSERV has been made since the original offering.

Return to the Section Table of Contents

Section II

II. Factors Leading to the Choice of Gopher for C401 and C402: Spring 1995

By the end of 1994, it was obvious that LISTSERV was far from the ideal medium by which to teach about chemical information sources. The effort to produce and load the files for instruction was very time consuming, and students were confused about how to retrieve the files.

By January 1995, the Indiana University Libraries had chosen gopher as the program through which to organize information about library services and collections. Thus, the chemical information courses offered in the Spring 1995 semester were adapted to gopher, and information about their availability was broadcast on CICOURSE and on CHMINF-L, the Chemical Information Sources Discussion List.[1]

Gopher offered several advantages over LISTSERV as a mechanism for delivery of instructional materials. First, Gopher's hierarchical arrangement of the course outline made it much easier for students to find the relevant lecture notes, exercises, etc. Second, the material was easily accessible at all times through the campus network. All UNIX machines included gopher clients, and all microcomputers in the public computer clusters had gopher clients. In addition, the variant of gopher used in the IU Libraries, GN, made the gopher files accessible to the burgeoning numbers of World-Wide Web browsers, such as Mosaic and Cello.

Although gopher access was a significant improvement over LISTSERV, there were still a number of drawbacks to its use. Gopher was fine for organizing class materials in a hierarchical fashion. Moving around in the hierarchy, however, required the students to retrace their paths and to remember where the instructor had placed a given document. Furthermore, gopher simply did not provide good access to the ever-increasing numbers of visual images and WWW resources that could enhance the teaching of chemical information.

Return to the Section Table of Contents

Section III

III. Factors Leading to the Shift to the Web

During the Spring 1995 semester, one of the course assignments for C402, Current Topics in Chemical Information, was to investigate chemistry resources on the Internet. Students had to choose the top ten resources and make presentations to the class that illustrated both more general sources and some that were specific to their areas of concentration in chemistry (analytical, biochemistry, organic, etc.). The enthusiasm with which the C402 students embraced the Web resources convinced me that all of the chemical information courses should ultimately be produced in HTML format. Consequently, during the final weeks of the Spring 1995 semester and the months prior to the start of the Fall 1995 semester, we set about creating the Web structure necessary to support all three courses.

There were a number of resources that were drawn on or adapted for use in the courses.

1. the existing C400 files from the LISTSERV and gopher efforts
2. the list of Internet Resources compiled by the author during 1994-95, "Some Chemistry Resources on the Internet" <="" a="">(2)
3. materials from the Clearinghouse for Chemical Information Instructional Materials (CCIIM)
4. training materials and database access provided via the Internet by commercial information companies and others.

By the start of the Fall 1995 semester, the Indiana University Bloomington campus had standardized on Netscape as the browser of choice. Netscape was available on all public microcomputers maintained by IU's University Computing Services, as well as on all PCs and Macintoshes in the IU Department of Chemistry's instructional computer facilities. Several of the microcomputers in the Chemistry Library also had Netscape, although the IU Libraries' public computers in general were not supporting Windows applications at the time. Furthermore, the campus had significantly increased the size and capacity of the modem pool, thereby enabling students who lived off campus to access the Web materials through a PPP connection. Supplementing the Netscape Web access was Lynx, the text-based Web browser found on all UNIX computers under University Computing Service control at IU.

Return to the Section Table of Contents

Section IV

IV. Formal Course Offerings in Chemical Information at Indiana University

The introductory course C400, Chemical Information Sources and Services, is a one semester-credit-hour course. Enrollments in the class, which is offered only in the fall semester, are large, typically in excess of 100. Most of the students are undergraduate chemistry majors, with a few graduate students from the specialized Masters program offered in conjunction with IU's School of Library and Information Science. Regrettably, there is no chemical information course devoted to chemistry graduate students at present, although all incoming chemistry graduate students in the C500, Introduction to Research, course are encouraged to attend several library lectures during their first semester.

The other one semester-hour classes have undergone considerable redefinition in the past few years. C401, Computer Sources for Chemical Information, as the name implies, concentrates exclusively on major computer-based tools. For example, during the Spring 1996 semester, three weeks were devoted to the Beilstein CrossFire system, and three weeks were spent on structure searching on the STN Registry File. Enrollments in that class tend to be in the range of 10-20 students. Finally, C402, Current Topics in Chemical Information, another one-hour class that can be repeated one time, offers the student the opportunity to concentrate on one or two topics which are explored in depth. During the Spring 1996 semester, the impact of the computer on the primary chemical literature (especially journals) and personal or in-house databases were the two themes of the course. This paper will concentrate on the Web applications for C400.

Return to the Section Table of Contents

Section V

V. The Basic Organizational Scheme for Chemical Information Sources on the Web

It was decided early on to use variants of the 15-week semester topic outline for C400 as the classification scheme into which the Internet components would be arranged. Thus, the material for the overview course (C400), the Internet guide to information resources in chemistry (CIS-IU), and the Clearinghouse for Chemical Information Instructional Materials (CCIIM) all have the following subject categories:

1. General introduction to chemical information sources
2. Directories and guides; Search strategies
3. General techniques of computer-based searching
4. Author, corporate source, and citation searching
5. Subject searching
6. Patent searching
7. Online chemical dictionaries and other sources for chemical compound searches (Chemical name and formula searching)
8. Chemical structure searching
9. Analytical (constitutional) chemistry searching
10. Physical property searching
11. Synthesis or reaction searching
12. Chemical safety or toxicology searching
13. Current awareness, background reading, document delivery
14. Science writing aids; The personal library
15. Miscellaneous information sources (history, biography, industry sources)
16. Teaching of chemistry.

Table 1. Subject Categories for C400, CIS-IU, and CCIIM

Maintaining a certain level of uniformity for classification of the material in the three separate products (C400, CIS-IU, and CHMINF-L) has made it easier to remember to draw on appropriate resources when preparing the lectures for the basic chemical information course, C400.

Return to the Section Table of Contents

Section VI

VI. Strategy for Integrating Web Resources into C400

The traditional course topics of the introductory C400 course were not modified significantly in either of the fall semesters when the network was utilized. The main difference was the elimination of most paper handouts for the students. Although this was a shock to some of them, they soon learned how to print out the most essential items. The other significant difference was the much larger array of potential sources that could be used as supplements to the basic material presented in the lectures. This was both a strength and a weakness, as some students became lost when they ventured off the local materials into Cyberspace. In such a rich environment of choices, with the world as the reserve room, it was difficult for some of them to decide which material was most germane to the course.

As with many other courses on the Internet, we developed a C400 Home Page from which the students could navigate to other areas of the course. However, the home page provided only one of several avenues to a needed document. For example, there is a "Problem Sets" section on the C400 home page that leads to another page where all of the problem sets (and selected answers) are listed. Additionally, the student could reach a given problem set from a hotlink in the Schedule of Classes.

The Schedule of Classes in fact became the central vehicle through which to tie together the focused lecture notes and the broader world of Internet resources relevant to the topic. Thus, for example, the section on patents is tied to the lecture notes through the lesson number and to the CIS-IU section on patents through the title of the section:

*Click number for lecture notes-->*6. Patents*<--Click here for CIS-IU*

Figure 1. Illustration of Hotlinks in the C400 Schedule of Classes

The other links to Internet resources from the Schedule of Classes were to the text of STN's Using CAS Databases on STN (and in future semesters will be to the new Using the CAS Registry File on STN). Thus,

at least a portion of the required reading for the course was taken from a remote site, the server at Chemical Abstracts Service. Note that in the Schedule of Classes, the STN lessons (indicated as CAS-1, Reg-1, etc.) are not assigned in the order in which they appear in the original work at CAS.

The bulk of the Internet connections in the C400 course are made through the lecture notes. Unfortunately, not all class lectures had completed Internet counterparts during the Fall 1995 semester, but it is hoped that at least some of the others will be finished prior to the start of the conference. As of this writing, 9 of the planned 15 modules had been completed. Another unfinished segment of the C400 site, which must probably await the advent of the 30-hour day for completion, is the Instructor Notes section.

Let us look at the lecture 7 notes on "Searching for Information on a Chemical Substance" in order to see how the Internet connections enhance the presentation of the material. Hyperlinked to the text of the lecture notes are materials produced locally at Indiana University as well as those reproduced from printed Chemical Abstracts Service documents (with their permission, of course!). In addition, Quick Reference Card materials located on the CAS server, chapters from the lessons prepared by Chemical Abstracts Service experts, a sample record downloaded from the Registry File, and the STN Database Summary Sheet for the Registry file, are among other links.

Return to the Section Table of Contents

Section VII

VII. Potential Problems with Hyperlinks

There is probably as much material *linked to* lesson 7 as there is actually *contained in* the text of the lecture notes. Imagine the plight of the student who might decide to print out all of lesson 7 and its hyperlinks in order to have a hardcopy more akin to the traditional textbook. It is doubtful that all of the bits and pieces of paper could be fitted into a correct sequence, especially if all of the links from the CAS documents were also printed out.

Another problem is that of disorientation in Cyberspace. The link from the subsection 7.2 of C400 lecture notes 7 entitled "Substance Searching Using Chemical Abstracts Service Registry Numbers" is actually to the CAS lesson 7 in *Using CAS Databases on STN*. At the very end of the CAS chapter is found "Return to Table of Contents," a hotlink to the cover page of *Using CAS Databases on STN*. This could be particularly troublesome because the CAS chapter does not indicate what source it is part of. Choosing that link causes the student to end up in a totally different document from that where he or she started. There is always a potential problem when one chooses not to link to the topmost page of the parent document at a remote location. Nevertheless, it was felt that linking directly to the individual CAS chapters gave a better fit to our lesson plans, allowing us to pick and choose relevant chapters.

The problem of disorientation is perhaps compounded by the inclusion of links from the C400 Schedule of Classes to the relevant sections of the Internet guide CIS-IU. The page on CIS-IU for "Chemical Name and Formula Searching" includes the section "Abbreviations of Chemical Compounds". Entries in that section lead to Web pages in Europe, where there are (fortunately in this case) few other avenues to be explored. However, the Patents section of CIS-IU links to a number of resources that can take the student into the backroads of the Internet.

In order to give a fixed orientation point from which to navigate, there should ideally be a bookmark in the browser the student is using to the homepage of the C400 course. Unfortunately for students using Netscape at IU public computer clusters outside the Chemistry Building, that is not possible, since Netscape is loaded from a server that does not allow the addition of new bookmarks. In the Chemistry computer classrooms, however, the bookmark for C400 is included.

Return to the Section Table of Contents

Section VIII

VIII. Fulltext Searching from the C400 Home Page

A fulltext search engine is presented to the student as the first item on the C400 Home Page. (3) The search engine identifies any relevant HTML document that begins with filename "400*.html" and resides on the www subdirectory of the cheminfo account at www.indiana.edu. For example, enter "chemical and business" as the fulltext search strategy from the C400 Home Page. One of the documents retrieved is actually:

Miscellaneous Information Sources: Biographies, Companies, Jobs, Teaching.

However, this shows up on the list of retrieved items as:

C400 Chemical Information Sources and Services Lecture 15.

That is because only the information in the <TITLE> section of the HTML coded document is listed in the results. Many of those titles for C400 documents are not the same as the actual titles seen when the documents are viewed with Netscape.

In order to find relevant information in the lecture 15 notes, the appropriate document would have to be scanned until section 15.5 "Information on Chemical Industries and Businesses" is reached. There is no easy way to jump to the section that caused the "hit" when this and most other Internet seach engines are used. Eventually a more detailed table of contents showing lesson subsections will be constructed to help navigate the C400 lecture notes.

The same type of fulltext search feature is included in CIS-IU. The search, "chemical and business" when run in CIS-IU would find more HTML documents (including those found in the previous search) because the CIS-IU fulltext search configuration file instructs the search engine to look at every document in the cheminfo/www subdirectory, not just the C400 documents.

Now try the search on CIS-IU by entering "chemical business". Dozens of items are retrieved! This illustrates one of the more confusing aspects of Web searching that confronts searchers today: lack of a uniform search algorithm. Students learn in C400 that on the STN system the default search logic for a multi-word subject search is **(W)**-- a proximity operator that is a more restrictive variant of the **AND** command. Thus, the STN system assumes that the words are to appear adjacent to one another in the order in which they were entered by the searcher. With the fulltext search engine on the IU system, the omission of the Boolean search operator defaults to an **OR** statement! Thus, any documents that have the word "chemical" are presented, as are any documents that contain the word "business". There is no requirement that both words appear in the same document, much less that they be adjacent.

Despite the peculiarities of the fulltext search option, it overcomes one of the serious obstacles of the gopher version of C400--the necessity to remember where a given document is located in the organizational scheme. The use of the fulltext search at least leads to the desired document (among others, perhaps) in the collection of HTML documents that constitute C400.

Return to the Section Table of Contents

Section IX

IX. Other Internet and Network Features Used in C400

Some of the exercises in C400 involve searching the STN LCA and LREG files. Students connect via telnet to STNC.CAS.ORG. The Internet connections were often quite slow and sometimes disconnected in the middle of the search. A much more serious problem is that the printouts do not show the actual commands input by the students, so in grading the assignments, we had to make educated guesses at times. Nevertheless, as the cheapest telecommunications connection to an online search vendor, the Internet will probably

remain our choice for online searching in the future.

The only other network feature used in C400 was e-mail. An e-mail distribution list of all students in the class made communication relatively easy. In fact, we have been using e-mail with students in the chemical information courses for several years now. All students at Indiana University can obtain free e-mail accounts, and few if any are hesitant to use them in communicating with the instructor. Some were even so bold as to contact the STN help desk by e-mail for assistance on the assignments.

Return to the Section Table of Contents

Section X

X. Conclusion

Judging by the number of people who skipped class during the Fall 1995 semester as compared to other years, the use of the Internet had an unexpected impact. Although unintentional, the introduction of a distance education capability apparently made some students feel that class attendance was less critical than in previous years. In looking ahead to the Fall 1996 C400 course, the Web material should enable us to change some of the activity that occurred in the past during the lecture hours. Thus, more demonstrations to illustrate some of the concepts or tools, some practice sessions with smaller groups in the computer classroom, and some discussion sections may replace the straight lecture format of previous years.

Even when the C400 class is not being taught, there is substantial usage of the material we have placed on the Web. A glance at the cheminfo statistics for the most recent month's usage confirms that. The "Total Transfers from Each Archive Section" reports which files are being used. This kind of feedback is useful and gratifying.

By and large, the Internet has empowered us to create tools and instructional materials that can be used anywhere by people who have an interest in chemical information. As the chemistry curriculum becomes more and more crowded, and as chemistry professors seek ways to allow undergraduate students to move into research, formal courses for chemical information instruction find themselves under assault in many academic institutions. At Indiana University, we are seeking ways to make available the course material to anyone who will take the time to learn it, whether in the context of a formal course or on their own. The Internet is a very significant tool in that quest, and we will continue to adapt our chemical information instructional efforts to take full advantage of it.

References

(1) Wiggins, Gary. "CHMINF-L: The Chemical Information Sources Discussion List," *Journal of the American Society for Information Science,* **1995 ,** *46*(8), 73-83.

(2) "Some Chemistry Resources on the Internet." The last version of this work was issued as number 10 (February 1995) and distributed to subscribers of CHMINF-L. It and earlier versions formed the basis of the Internet Chemistry Resources list at Rensselaer Polytech.

(3) Fulltext Users' Guide at Indiana University

(4) For other Internet courses relevant to chemical information, click here.

Return to the Section Table of Contents

Return to CHEMCONF '96