Internet Resources for the Humanities and Social Sciences

Catherine Sealey, Chadwyck-Healey Ltd.

The easiest and quickest way to handle large amounts of data is to put it on the web. Large databases on the Internet empower researchers to get the information they need quickly and easily, regardless of their physical location. In addition the Internet provides a high level of interaction between people and information, allowing manipulation of the data. As a result, the Internet as a research tool is growing at an unparalleled rate and is revolutionising the manner in which academics work.

Chadwyck-Healey is developing a number of major websites which bring together many different databases in the same discipline or subject area to form a home for that discipline on the Web. These databases have proved to be extremely popular with researchers in the humanities and social sciences. They combine full-text materials with bibliographies and reference works and links to related information elsewhere on the Internet. They make all the information accessible through a simple interface.

All of Chadwyck-Healey's Web databases are HGML or SGML coded allowing each database to be fully searchable and uses standard interfaces familiar to everyone; Netscape Navigator, Microsoft Internet Explorer so librarians do not have to instruct users on how to use them. The Databases are also extremely accurate. The data is typed in by hand twice by two different operators, and a comparison program highlights any discrepancies for review. All illustrations are scanned and stored as images and are linked to the appropriate point in the text by the coding.
Consequently the cost of these databases is high, but the price is less than a University library would have to pay for the same works in book form.

(Slide 1) Who is Chadwyck-Healey Ltd ?

Chadwyck-Healey started off as a Microfilm company in 1973, specialising in large Microfilm products for the international library market.
The Microfilm markets expanded in the 1970's and 1980's and with the introduction of CD-ROM technology in 1986, the company embraced the new technology and became one of the first companies to offer databases on CD-ROM.
In 1996 Chadwyck-Healey launched it's first Internet service, PCI Web. The company now offers three major web sites: Literature Online, Deutsche Literature Online and Know UK.

The company still provides hundreds of Microfilm titles and now has about eighty CD-ROM titles specialising in social sciences business and humanities.

The company started in 1973 with about 10 people. This has now grown to 130 people working at the Head Office in Cambridge, we also have a sister company in the USA and a company in Spain.

(Slide 2) Periodicals Contents Index (PCI)

The Periodicals Content Index database was the first database which Chadwyck-Healey converted on to the web. It is the largest single Chadwyck-Healey publication with currently 8.9 million records. The collection consists of 14 CD ROMS.

The Periodicals Contents Index is a fast efficient index to the contents pages of more than 2,257 journals from their inception to 1990. Where other electronics indexes tend to cover only the last twenty years or so. PCI indexes journals from the eighteenth century to the present day. The Periodicals Contents Index is not a "current contents" project but rather a retrospective database.

PCI web catalogues every article in thousands of scholarly journals. It indexes complete runs of journals, some back to 1770. It covers every area of the humanities and social sciences, together with journals of general interest. Its scope is international and is available in around thirty different languages.

PCI web is an ongoing project.
An experienced team of researchers and editors are working with current subscribers to identify more journals for inclusion.
New segments are added every two months, so that the index grows by more than one million records each year. PCI web will eventually encompass, 3,500 journals and 15 million articles.

The database will open up a library's periodicals collection and enables user to fully exploit these valuable resources. PCI Web is already used in National libraries, University libraries and Public libraries throughout the world. Even small libraries without a huge range of journals can use PCI web to point researchers in the right direction.

It cuts down the research time and allows rapid initial research to a complete set of journals making sure all relevant literature is taken in to account.
PCI Web is available on annual subscription in two editions.
Complete, with coverage from 1770 to 1990/91; and subset, with coverage from 1960/61 to 1990/91.

(SERARCH DEMONSTRATION)
PCI Web allows you to search by journal subject, general keyword, article title keyword, author, heading, journal title or language. You may search through the whole of PCI web or select a range of dates. You may choose to search in all records, exclude book reviews or search only in reviews.

LITERATURE ONLINE (Slide 3)

Literature Online is a vast electronic library which combines almost a quarter of a million texts of poems, plays and novels including every know poem written in England before 1900, 4000 complete plays; 290 works of prose fiction; and 21

versions of the Bible in fully searchable SGML format with reference works, bibliographies and catalogues on a single site, and provides hypertexts links and indexing for relevant resources on other web-sites.

Nine full text databases originally on CD-ROM are brought together in an online context, making them fully cross searchable. A huge literary site has been created by adding secondary resources such as: dictionaries and bibliographical and biographical databases. Resources used are:
The bibliography of American Literature, Webster's Third New International Dictionary and ABELL : The Annual Bibliography of English Literature and Language 1920 to the present.

Literature Online is of interest to students of English and American literature. The subject matter contains English literary works; Early English Prose fiction, 1500-1700. Edition and Adaptations of Shakespear 1591-1911, Eighteenth Century Fiction 1700-1780, English Drama 1250-1900, English Poetry 1100-1900 and Modern Poetry 1972-1997 and also specialist subject areas; African-American Poetry (1750-1900) and American Poetry (1600 -1900).

The database is held together by a Master Index which lists down to poem title level every work in LION, more than 200,000 entries. Clicking on a title takes you directly to it whatever database it is in.. This give LION an organisation and structure which differentiates it from the Internet.

Also accessible through the Master Index are thousands of other literary texts created by academics or their institutions on freely accessible web sites. In this way a further 20,000 texts together with discussion groups, literary sites and library catalogues are embraced by LION, which will continue to grow more and more.
Book buying facilities are included. So that the user can book up the price and availability of a title, order by email and pay by credit card.

Literature Online is a growing database, new texts are regularly converted to electronic form by Chadwyck-Healey and licences for reference works are negotiated with other publishers. Resources at other websites are continually being identified, linked and added to the Master Index.

Further additions to Literature Online planned this year include;
American Poetry 2, covering Twentieth Century American Poetry from 1919 to the present.
Essays on the American Novel, a collection of essays, published by Cambridge University Press containing commentaries on the major work of American fiction.
The works of W.B. Yates, a fully searchable collection of Yeat's principle works, his poems, plays, fiction, critical and autobiographical writings.
Nineteenth Century Fiction , providing the first opportunity to cross search the full text of the major nineteenth Century Literary form containing more than 200 novels of English prose fiction such as Frankenstein and Pride and Prejudice.

(SEARCH DEMONSTRATION)

(Slide 4) International Index to Music Periodicals (IIMP)

The International Index to Music Periodicals indexes over 350 international music periodicals from more than twenty countries. IIMP is the only electronic index with extracts to current music periodicals. New journals are added with the advice of the IIMP Advisory board.

IIMP currently contains over 160,000 records including records from twenty-eight journals from their inception through 1995, with complete runs of the
Journal of American Musicological Society and Notes and records from 100 journals from 1991-1995. IIMP is both current and retrospective with over 30,000 records dating from 1996 to the present and other records dating as far back as 1874. IIMP is growing daily with many new features to be added in 1998, including 100,000 more retrospective records.

The International Index to Music Periodicals is updated quarterly. The articles in the database are indexed by 9 terms; author, title, publication, publisher, subject heading, date of publication, language, article type and country.

We are also producing a full-text version of "The International Index to Music Periodicals" which will be available in May this year featuring initially the full text of current (1996 forward) articles from 26 of the most important music titles
including; American Record Guide and International journal of Music Education. Monthly updates to the database are planned.

IIMP full text will also include links to the full text of indexed e-journals. Retrospective addition of full text to 1996 citations will begin in Autumn 1998.

(Slide 5) Guardian on the Web

We will be soon be launching "Guardian on the Web" planned release is expected in the summer of this year.

The Guardian is the newspaper of choice for the scholarly community and is widely read and consulted in Europe.

The Guardian is currently available on archival CD-ROMS containing the full text of the newspaper from 1990 to 1997, with two years on each disc.

The web version will create a fully searchable full text database from 1990 with daily updates.

(Slide 6) Patrologia Latina Database

The Patrologia Latina Database is the full-text electronic version of the Patrologia Curus Completus; Serie Latina published by Jacques Paul Migne 1844-1857, including all prefatory material, original texts, critical apparatus, indexes and illustrations.

The database consists of the 221 volumes on five discs and is the most complete and authoritative collection of writing of the Church fathers such as Augustine. Jerome, Hilary and Bede.

It features over 1000 writers from Terullian in AD200 to Poe Innocent III in AD 1216 and spans over 1,000 years of Church History.

The World wide Web makes this massive resource more widely available. Libraries can subscribe to it annually, at a fraction of the cost of the CD-ROM edition.

Individual words are searchable in Greek or Latin anywhere in the text or title of a work. It is possible to serach works of a single author or search across all at the same time. You can truncate your search terms to find all forms of a word using proximity operators, you can serarch for words near to each other. Any search term can be combined with any other Boolean operators.

The Patrologia Latina Database has made an enormous difference to the work of scholars of patristics, theology, medieval history and Latin. In many other disciplines in the humanities, from philosophy and medieval languages to the history of art and the history of medicine, it is making significant contribution to research.

To summarise producing literature on the Internet gives access to a single database without the problems associated with networking CD-ROMS such as worrying about the software, the number of drives needed or having to take a CD-ROM off a shelf to install it. The problem of regularly updating information is also solved.
It is true that the customer does not have ownership of the database, but for many Libraries access has become more of an important issue.

Web delivery also offers control of access and the monitering of usage in a way that is impossible with CD-ROM, offering the libraries greater control over their information.

Publishing on the Internet is cost effective and allows libraries to take advantage of the webs linking ability and transforms the potential of a database. LION for example is almost a virtual library of literature. The Internet takes the whole publicity process a stage further.

Chadwyck-Healey Ltd
The Quorum
Barnwell Road
Cambridge
CB5 8SW
Tel: 0044 1223 271253
Fax: 0044 1223 215514

Contact: Catherine Sealey, Sales Executive
email: sealey@chadwyck.co.uk