

An International Digital Library for Distance Learning

*Research issues for the UK Open University's
International Centre for Distance Learning (ICDL)*

Nick Farnes
International Centre for Distance Education
The Open University, UK

Abstract

The paper describes the International Centre for Distance Learning which is located in the Open University's Institute of Educational Technology. An essential resource of the Centre built up over 15 years is its digital library which includes information on 1,200 distance education institutions, 31,000 courses and 10,000 abstracts from the literature of distance learning. It is available on the WorldWideWeb and receives over 10,000 hits a day from 25,000 users in over 130 countries. Services promote research collaboration, the exchange of experience, course development and learning opportunities. This paper outlines the research issues for ICDL, set in the context of the Open University and current developments in its teaching system, which is undergoing fundamental change. It is hoped that ICDL's digital library and its research programme, will assist in developing new paradigms for distance learning and through its contribution to the OU's MA in Open and Distance Education provide a test-bed with implications for other discipline areas.

What is ICDL

ICDL promotes international research and collaboration by providing information based on its library and databases; it reaches other audiences through publications, including ICDL Update, compilations and reviews; teaches both face-to-face and online in the UK and abroad, undertakes international consultancy; and conducts its own research. An essential knowledge resource of the Centre built up over 15 years is its distance education library and databases:

Library - literature

The mainly paper based library is the largest collection of literature on distance learning in the world. There are over 15,000 items - books, journals, journal offprints, conference proceedings and papers, research reports, surveys, dissertations, and newsletters - so-called grey literature. ICDL aims to obtain copies of all new books and monographs published in English on distance education, open learning and related fields. It subscribes to all the specialist distance learning journals and receives newsletters published by institutions and organisations involved in open and distance learning. Most of the collection is in English, but there are significant numbers of documents in French, German, Spanish and Chinese.

Library - institutions

It also has the largest repository of information about institutions across the world including prospectuses, calendars and course handbooks. Included are institutions dedicated to distance teaching, open universities, institutions which have distance teaching departments and conventional, eg. dual mode; or have programmes or courses taught at a distance. Coverage includes all education and training levels.

Databases

Information is provided online through the WorldWideWeb (<http://www-icdl.open.ac.uk/>)

- literature on distance learning - 10,000 entries
- distance teaching institutions - 1,200 entries
- distance taught courses and programmes - 31,000 entries

The databases are searchable by subject, author, country, region, date of publication, subject, educational level, institution and in a variety of other ways. Every item listed (and abstracted) in the database is also held in the library collection.

To support the teaching on OU online courses and the MA in Open and Distance Education restricted access sites providing a database of selected distance learning literature with links to full text versions of papers. Currently these resources are classified under 60 main headings with up to 10 items under each. ICDL also contributed to a database and is a partner in the World Bank Global Distance Education network, which has been established to provide core and regional information online for development policy makers and project managers (<http://wbweb4.worldbank.org/disted/>).

In order to extend the knowledge base ICDL coordinated the ManageLearn SOCRATES project with partners from institutions in Spain, Portugal, Germany and France. The project established a database of literature on open and distance learning for higher education in Europe (Harry, 1997a). Also a prototype multi-lingual thesaurus of key words was developed which provides a multi-lingual search facility (<http://www4.open.ac.uk/Managelearn/>). Links to the distance education databases of other partners have been set up.

Who uses ICDL

The largest group of users are those online. Currently ICDL receives up to 10,000 hits a day from 25,000 users in over 130 countries which adds up to 300,000 hits per month. Approximately 90% of requests are for institution and course information, with 10% using the literature database. The build up since the databases went onto the WorldWideWeb has been extraordinary. Prior to November 1966 the databases were accessible on the Internet but through an interface involving passwords and had an average daily use of around 300 from 150 registered users.

Where do users come from? - The main countries from which users log on from are shown in the table below:

Table 1 Top 20 countries where log-ons originated (figures for one month)

<u>Country</u>	<u>Total hits</u>	<u>% of total</u>
US	85591	27.56
UK	33414	10.76
Canada	16506	5.31
Australia	6246	2.01
Singapore	4995	1.61
Malaysia	4958	1.60
Japan	3573	1.15
Germany	3405	1.10
Netherlands	2927	0.94
Greece	2758	0.89
New Zealand	2504	0.81
Brazil	2204	0.71
India	2144	0.69
South Africa	2056	0.66
Hong Kong	1922	0.62
Italy	1825	0.59
Sweden	1776	0.57
Belgium	1669	0.54
Spain	1629	0.52
Ireland	1406	0.45
		(59.08%)

Regarding the European use of the databases, the hits by each country, grouped by east and west are given below.

Table 2 Number of hits per month from East and West European countries

<u>E Europe</u>	<u>Hits</u>	<u>W Europe</u>	<u>Hits</u>
Czech Rep	686	UK	33414
Russia	556	Germany	3405
Croatia	372	Netherlands	2927
Poland	361	Greece	2758
Hungary	317	Italy	1825
Latvia	190	Sweden	1776
Romania	181	Belgium	1669
Yugoslavia	154	Spain	1629
Bulgaria	114	Ireland	1406
Slovakia	108	France	1345
Estonia	84	Norway	1272
Bosnia	83	Switzerland	1261
Slovenia	58	Denmark	830
Lithuania	41	Austria	690
Ukraine	22	Finland	666
Albania	6	Portugal	559
		Cyprus	362
		Malta	311
		Luxembourg	137

Summary figures for western Europe excluding the UK, eastern Europe and the UK are given in the table below as well as the US figures for comparison.

Table 3 Summary of hits for East and West Europe, the UK and US

	<u>Hits</u>	<u>% of total</u>
W Eur (ex UK)	24828	7.99
E Europe	3333	1.07
UK	33414	10.76
Total Europe	61575	19.83
Total US	85591	27.56

While the content of ICDL's digital libraries are distance learning and its methods involve collaboration and networks, it is not a distance learning system in the same sense as an open university or other programmes of distance education. This is mainly because it does not provide teachers or a tutorial system, nor does it accredit the learning it promotes. However, it does operate key functions which promote and support distance learning. It collects and creates knowledge resources, which are organised in a structured and systematic form; makes these resources accessible at a distance and delivers to individuals wherever they are; provides tools to assist users in identifying their needs and engaging with the resources; furthermore it is open to all and not restricted to fixed times or dates.

ICDL is used mainly by:

- 1) distance learning researchers, scholars and students engaged in the field
- 2) distance learning specialists who are developing opportunities for others to learn
- 3) potential students or their advisors seeking appropriate distance learning opportunities.

It promotes collaboration between creators and users of knowledge and is part of a global knowledge network.

Research

ICDL not only supports and disseminates the work of thousands of researchers worldwide and promotes their learning, it also conducts its own research on comparative distance learning, knowledge networks and pedagogy. The research issues focus on:

- Knowledge base
 - *the field of distance learning*, its key words and concepts, structure, boundaries, research agendas, authors, and changes; on the institutions and courses, their nature and subjects, comparative research, trends, reviews.
- Users and uses
 - *ICDL users*, who they are, where from, what they want information for, their role in the field of distance learning.
 - *how ICDL is used*, what information users access, their pattern and frequency of use, growth, social geography of distance learning, who is accessing whose information, user and producer overlaps, international cooperation, globalisation.
- Evaluation
 - *the impact of knowledge about distance learning*, how used, what effect, benefits - further research, quality of research, use to produce more information resources, use in distance learning, impact of ICDL, ICDL's role in field and as part of a global knowledge network.

Knowledge base

Attempts have been made to represent the knowledge structure underlying distance learning as a discipline and its boundaries with other related disciplines. The classification system used by ICDL is one example from a library tradition and there are other classification systems applied to the field. Changes over time to the topics of research have been documented (Harry, 1997b; Panda, 1995) and studies involving citation analysis have been carried out focussing on geography (Calvert, 1995) and over time Bunker 1998).

With digital libraries the potential for computer analysis of their contents can be exploited. There is a long tradition of research into information retrieval (see Sparck Jones and Willett, 1997) and knowledge modelling (eg. Doyle, 1962; Schvaneveldt, 1990). To address the general question about the nature of the field of distance learning various analytical tools and software can be applied, these include:

- Concept analysis: word frequencies, keywords, co-occurrences, clusters, structures
- Author analysis: number of entries, productivity, key figures, co-authors, citation analysis, active period
- Document mapping: proximity, similarity, profile matching, linkages

The procedures for analysing the word frequencies, establishing key words and concepts are well documented and software is available (CTI, 1999; Berber Sardinha, 1996). Measures of proximity or association between the main concepts include the frequency of their co-occurrence. This data can be used for creating matrices showing the strength of association between pairs of concepts and provide the input for network and modelling software (Schvaneveldt, 1990; Chen, 1998). The resulting structures can give an overview of the conceptual structure of the field.

Similar work has been carried out on citation analysis in other disciplines to show networks of referral, influence patterns and the social structure of a field of knowledge (Reid, 1997; Chen and Carr, 1998).

The extent that documents share key words and concepts can also be used as a measure of their proximity and in a similar way the strength of association between pairs of documents can be used to give an overall structure to the main writing in a field (Fowler and Dearholt, 1990).

A collaborative project is underway involving ICDL, the Open University's Knowledge Media Institute and Brunel University's Department of Information Systems and Computing. The intention is to apply knowledge modelling expertise developed by the collaborators to entries in the ICDL literature database to model the domain of distance learning and install navigation tools. The analysis will also be applied to H801 a post-graduate distance learning course - Foundations in Open and Distance Education. The research aims to assess the potential of knowledge modelling for:

1. Monitoring, reflecting on and facilitating the advancement of the subject area of distance learning
2. Curriculum and pedagogic developments for courses in this domain
3. Course development in other subject domains

The pedagogic implications of this work for distance learning are considered in the final section of this paper.

Users and uses

Analysis of user searching and browsing can be carried out from computer logs. Collection and analysis of user defined search words gives an indication of their perspectives on the field, and their searching by author

suggests the most popular authors which can be compared with the authors of frequently visited entries. Whether the most frequently visited entries are related to the most productive authors is also being investigated. A measure of the number of visits per entry shows a distribution from the most popular - visited 267 times in 1998, through to the majority of entries visited less than 10 times. Work on users interests in particular institutions and courses is also underway.

Measures of association or proximity between entries can be established by analysing the pairing or clustering of entries visited. In addition to frequently visited pages, their nearest neighbours to which users often go onto visit can be assessed. Matrices showing the frequency that pairs of entries are visited can be constructed and used as the basis for modelling the structure of the key entries, and provide an overview of users' perceptions of the core writings in the field.

As an indicator of ICDL's contribution to international collaboration, analysis of records can show the extent that users are requesting information about their own institution - as is the case in some African institutions where the database entries are the most comprehensive available - or about institutions elsewhere in the same country or from other countries. We expect the patterns to depend on the particular countries concerned and show whether ICDL is being used to access knowledge internationally or if the field is largely nationally and regionally biased as was shown by Calvert's (1995) study of journal citations where authors tended to cite others in the same country or region.

Evaluation

The main elements for the analysis of the contents and use of a digital library are represented in table 4 below. The main actors are authors and users. Authors apply the words and concepts of their discipline to construct documents which are entered into databases, the entries include author names and over time represent the productivity of each author in the field. Users access documents using search words and author names, and some documents and authors are more popular than others.

Table 4 Main elements and main actors

Element	Authors	Users
Key words/concepts	word frequencies in entries	frequently used search words
Names	productivity	popularity of authors
Documents	number per author	frequency of visits

To model the contents of a digital library it is necessary to use a measure of association between the items under each of the main elements. Generally used measures are based on either author or user behaviour (ie. what authors write and how and what users access) give values of co-occurrence and proximity between pairs of items. Table 5 shows the main elements and measures that can be used to draw up association matrices which provide the input to networking and modelling software.

Table 5 Measures of association between the main elements

Element	Authors	Users
Key words/concepts	co-occurrence, pairings, linkages, models	search word sequences and pairings, clusters
Names	citations, co-authoring, networks	authors search sequences and pairs
Documents	key word matches, proximity, clusters, mapping	next visited, pairs, clusters

These analyses can be applied to the whole database and user behaviour over a specified period. Also the database can be segmented to show the changes over time in the main concepts, authors and documents and in their relationships. In this way the development of the field can be plotted and areas of new and rapid developments identified (eg. Web-based teaching since 1994).

One way of looking at research projects and publications is to regard previous documentation as inputs to the process of knowledge creation and new documents as the outputs. Of course there are other inputs (eg. discussions with colleagues, research data, observations and experience etc) but citation analysis is one way of considering the inputs and outputs of knowledge creation. However, the role of a particular library in enabling the connections is not clear cut and authors may have found out about and obtained documents from

many sources. User surveys can gain some information regarding the contribution of a particular digital library.

Research on the utilisation of information has shown that this is a complex process. It is rarely a matter of simply finding information and applying it. A distinction has been drawn between 'instrumental' (immediate and directly observable) and 'conceptual' (delayed and diffused impact and less observable) and this might reflect the difference between 'information' and 'knowledge'. All forms of information are not the same and there is a need to specify typologies of information and uses and the conditions and circumstances under which various types of information can be employed for different types of problems (Rich, 1991). In other words to take a contextual and contingency approach in order to assess the use of information and the impact of what has been learnt (Farnes et al, 1994). Knowledge does not exist in isolation, its use depends on contextualisation both upstream (ie. where it comes from) and downstream (what it is to be applied to). Consequently research and evaluation studies on the use of ICDL's knowledge resources need to relate who users are, their circumstances, and the task - and recognise that any action taken will be subject to multiple factors and not determined simply by a single item of information.

The role of digital library support for distance learning can be assessed through content analysis of transcripts of online collaborative learning and students assignments. Also end of course feedback surveys can ask explicitly about the use and usefulness of digital library services (Farnes, 1993; Jelfs, 1998). In general the benefits of such services depend not just on the nature and scale of the online resources available, but also on the pedagogic context of students' learning. If the library is simply an 'add-on' feature rather than integral to the mode of study, then its usefulness will obviously be less. However, this simple point is often overlooked in the haste with which distance teachers introduce digital resources to courses, without establishing an adequate framework for students to engage with the resources (and each other). The pedagogic implications of digital libraries are considered in the next section of this paper.

Distance Learning

A comprehensive review of libraries for distance learning concluded that two universal themes emerged from the 518 works cited in the bibliography - 'the recognition that distance learners in every country require access to appropriate library and information resources and the genuine desire of librarians to assist distance learners in obtaining access' (Slade and Kascus, 1996). However, the review suffers from an unintended bias, in that all the documents are written by those with an interest in library services for distance learners. The strong case for providing distance learners with all the materials and support needed for successful independent study without library facilities does not come through. For example, the teaching system for the world's pre-eminent open university was originally conceived explicitly to exclude the requirement for library support of any kind.

The British Open University, established in 1969, was originally designed as a mixed media teaching system, using the mass media - broadcast TV and radio; structured print study guides; set books and a student support system of tutors, counsellors, local study centres and short residential schools. It depended on huge upfront investment of academic and professional staff, working in course teams over a period of 2-3 years. The University created, assembled and delivered all the teaching materials students needed for their studies, except for a limited number of published set books which they were expected to purchase. There was no need for students to have access to libraries. Furthermore the students were distributed throughout the UK and their access to public libraries could not be guaranteed, and only a few would have been able to reach an academic library. The OU did not follow the smaller scale Australian dual mode universities which set up their own mail order library services.

This model of distance education has been described as second generation involving the industrialisation of education with mass production, division of labour and economies of scale. Second generation represents a step change in cost, scale and complexity from first generation which tends to operate on a small scale with courses produced and taught by a single teacher, using a single medium (eg, print, and sometimes radio). We are currently witnessing the forging of a third generation of distance teaching based on electronic media and a different teaching paradigm (Bates, 1991; Nipper 1989).

OU distance learning courses can be seen as representing a knowledge domain constructed from the collaborative efforts of the course team drawing on their knowledge and other knowledge resources. Furthermore students' knowledge is the product of interaction with the course, other students and their tutors (and others) and is represented in their assignments, projects and examinations.

Increasingly the teaching model is moving towards resource based learning within a pedagogic framework and the use of online conferencing (and assignments) to enable students to engage with resources (including each other),

construct knowledge and produce assignments, with electronic submission and feedback. The knowledge embodied in second generation structured course materials is more explicit than that contained in resource material and arising from students' interactions (eg. collaborative learning). Moving towards a less firm knowledge base, makes it more difficult to assess the quality of learning and assignments. More flexible, resource based and student centred approaches also require appropriate structures. The success of the OU system would be undermined by open curricula, unstructured pedagogy, subjective assessment. Structure is necessary in the relationship between knowledge resources, acquisition, construction and production/reproduction. Knowledge modelling techniques applied to resources, students' interactions and assignments might assist in providing structure and the promotion of effective learning.

Digital libraries and new technology generally are *not* distance learning systems. They can provide a means of facilitating learning but do not provide the value added by a learning system. Value added functions include:

- Entry, admission, registration, a learning contract of some kind
- Definition of the curriculum, knowledge and skills, what is included
- Series of tasks, instructions, learning activities
- Motivation, progression and commitment
- Feedback, support and guidance
- Accreditation, certification

To organise these services a learning system with infrastructure, management, operational logistics, resources and scale is required. The idea that all that is needed for resources can be loaded onto the Web and learners can simply access this material is attractive to policy makers. However, it fails to recognise that structured learning requires more than clicking on Web pages. In particular there needs to be a human element to provide support, feedback and guidance - and most important to judge and award recognition of achievement. In essence a learning system must provide an appropriate degree of structure. With too little structure sustained progress and success are more difficult, too much and rigidities and inflexibilities create barriers for participation, increase possibilities of failing to meet deadlines and induce frustration in learners.

The work on the analysis of the content and structure of a field of knowledge is important for monitoring the development of the discipline and its coverage in a digital library. The resulting measures of association and structures can be used to help users navigate the field, either to find related documents, or to explore related concepts. Incorporated into a distance learning system a digital library can have a central or supporting role. The crucial issue is to provide appropriate structure for the curriculum and pedagogy for successful distance learning.

References

- Bates A W (1991) Third Generation Distance Education: The Challenge of New Technology, *Research in Distance Education*, April, pp 10-15
- Berber Sardinha A P (1996) Review: WordSmith Tools, Computers and Texts, no 12, July. Also available at <http://infor.ox.ac.uk/ctitext/publish/comtxt/ct12/sardinha.html>
- Bunker E (1998) Gaining perspective for the future of distance education from early leaders, *American Journal of Distance Education*, vol 12, no 2.
- Bunker E (1998) Speakers in the ICCE/ICDE conferences, part II: Citations, Ed D dissertation - An historical analysis of a distance education forum: the International Council for Distance Education World Conference Proceedings, 1938-1995, Penn State University, Pennsylvania.
- Calvert J (1995) Mapping Knowledge in Distance Education, proceedings of ICDE Conference, Birmingham, The Open University, Milton Keynes.
- Chen C (1998) Bridging the Gap: the use of Pathfinder Networks in visual navigation, *Journal of Visual Languages and Computing*, 9, pp 267-286.
- Chen C and Carr L (1998) Trailblazing the literature of hypertext: author co-citation analysis (1989-1998), <http://www.brunel.ac.uk/~cssrccc2/>
- CTI Textual Studies: Guide to Digital Resources (1998), <http://info.ox.ac.uk/ctitext/resguide/resources/index.html>
- Doyle L B (1962) Indexing and Abstracting by Association, reprinted in Sparck Jones and Willett (1997).
- Ellis D (1996) *Progress and Problems in Information Retrieval*, Library Association Publishing, London.
- Farnes N C (1993) Evaluation of an Online Education and Training Course, Report for JANUS/DELTA evaluation, available as SRC Report, IET, The Open University, Milton Keynes.
- Farnes N C (1998) A Distance Education Knowledge Network and Global Learning System - new developments at the UK Open University's International Centre for Distance Learning, paper presented at International Conference on Collaborative and Networked Learning, IGNOU, India.
- Farnes N, Woodley A and Környei I (1994) How Distance Learning assists in the Transition towards a Market Economy, proceedings of the European Distance Education Network (EDEN) Conference, Tallinn, Estonia, EDEN, Milton Keynes.

- Fowler R H and Dearholt D W (1990) Information Retrieval Using Pathfinder Networks, Chapter 12 in Schvaneveldt.
- Harry K (Ed) (1997a) Open and Distance Learning: a selective bibliography for higher education institutions in Europe, ManageLearn SOCRATES project, ICDL
- Harry K (1997b) Research Reviews and Agendas, Section 5, *Researching the Literature on Open and Distance Education*, Block 3, H801, *Foundations of Open and Distance Education*, The Open University.
- Jelfs A (1998) Evaluation of H801 - Foundations of Open and Distance Education, internal report, IET, The Open University, Milton Keynes.
- John M (1998) ManageLearn SOCRATES project, 2nd year report, ICDL, Milton Keynes
- Nipper S (1989) Third Generation Distance Learning and Computer Conferencing, in Mason R and Kaye A (Eds) *Mindweave: Communication, Computers and Distance Education*, Pergamon, Oxford.
- Panda S K (1995) Research and Development in Open and Distance Education, in Singh B (Ed) *New Horizons in Distance Education*, Uppal Publishing House, India.
- Reid E O (1997) Evolution of a body of knowledge: an analysis of terrorism research, *Information Processing and Management*, vol 33, no 1 pp 91-106.
- Rich R F (1991) Knowledge Creation, Diffusion and Utilization: perspectives of the founding editor of Knowledge. *Knowledge*, vol 12 no 3 pp 319-337.
- Schvaneveldt R W (Ed) (1990) *Pathfinder Associative Networks: studies in knowledge organization*, Ablex Publishing, New Jersey.
- Slade A L and Kascus M A (1996) *Library Services for Off-Campus and Distance Education, the second annotated bibliography*, Libraries Unlimited, Colorado.
- Sparck Jones K and Willett P (Eds) (1997) *Readings in Information Retrieval*, Morgan Kaufman, San Francisco.