

How did university departments interweave the Web: A study of connectivity and underlying factors

Chaomei Chen^{a,*}, Julian Newman^b, Rhona Newman^c, Roy Rada^d

^a*Department of Information Systems and Computing, Brunel University, Uxbridge UB8 3PH, UK*

^b*Department of Computer Studies, Glasgow Caledonian University, Glasgow G4 0BA, UK*

^c*Department of Sociology, University of Ulster at Jordanstown, Newtownabbey BT37 0QB, UK*

^d*School of Electronic Engineering and Computer Science, Washington State University, Pullman, USA*

Received 1 February 1997; received in revised form 1 May 1997; accepted 1 November 1997

Abstract

This paper presents two studies of the use of the WWW in Scottish universities and American land-grant universities. First, we investigated the relationship between the organisational profile of a university department in Scotland and its structural connectivity on the WWW. A Spearman rank order correlation analysis revealed a number of strong correlation relationships between structural connectivity measures and the organisational profile based on research assessment exercise ratings, teaching quality assessments, student–staff ratios and funding levels. Linkage patterns from 13 Scottish academic sites to commercial sites in Britain and America highlighted the impact of culture and the appropriateness of information technologies on the acceptance of the WWW. The second study is a content survey of WWW-based education activities in American land-grant universities to investigate successful applications of these enabling techniques in education. The two studies together highlighted cultural, political and technological interactions in the use of the WWW. © 1998 Elsevier Science B.V. All rights reserved

Keywords: WWW; Connectivity analysis; Hypertext references; Scientific networks; Content analysis

1. Introduction

The World-Wide Web (WWW) is the first realisation of a global and distributed hypertext. In 1945, Vannevar Bush envisaged Memex, a pioneering vision of hypertext,

* Corresponding author.

as “a sort of mechanised private file and library” and “a device in which an individual stores his books, records, and communications, and which is mechanised so that it may be consulted with exceeding speed and flexibility”. Ted Nelson conceived hypertext as a universal repository for everything that has ever been written and everything is deeply intermingled in the hypertext—a distributed and global hypertext.

The introduction of graphical browsers such as Mosaic[™] and Netscape Navigator[™] tremendously contributed to the globalisation of the WWW. The WWW not only provides an enormous increase in the information accessible from the user’s desktop across all the major platforms, but also enables users to choose information services distributed across different countries.

The globalisation of information technology may interact with cultural bases in terms of technology transfer, innovation diffusion and user acceptance [1]. The WWW has rapidly achieved a global acceptance and adoption by users in different cultures. The WWW integrates a set of enabling technologies for people to get to know about each other and to form a virtual community. One would expect that, on the one hand, the enabling technologies provided by the WWW may forge effective knowledge diffusion and technology transfer as users are less and less constrained by geographical boundaries and different time zones; on the other hand, the appropriateness and acceptance of the technology itself may be strongly shaped by local cultural and economic environment.

Many demographic surveys explored the breadth and depth of the globalisation, such as WWW User Surveys at Georgia Institute of Technology [2] and University and College Home Pages Survey [3]. On the other hand, it is largely unknown how the vast amount of information is interweaved on the WWW across different companies, universities and countries. What is the nature of such interconnectivity? What can we learn from these connectivity patterns about the impact of culture on the acceptance and perceived appropriateness of technology?

In this paper, we focus on how universities fit themselves into the new dimension of an information society. We analyse the connectivity of computer science departmental websites across 13 universities in Scotland. We trace several types of destinations of hyperlinks to identify important characteristics of social and scientific use of the WWW. We obtain various structural statistics such as incoming hyperlinks and outgoing links associated with each website. We also build the profile of each institution on the basis of research and teaching resources and performance assessments associated with the institution. In addition, we explore the nature of linkage patterns from Scottish universities to commercial sites in Britain and America. A survey of WWW-based education activities in American land grant universities is also presented to highlight the use of the WWW in different cultural settings.

1.1. Scientific networks and cross-institution collaboration

Latour [4] characterises the nature of scientific research as a network of activity, in which each individual investigation is built on peer researchers’ work as well as on recognised work. A network of scientists who work at different institutions but

who regularly exchange working papers and technical memoranda, attend the same conferences and seminars, and build upon one another's work, is known as an 'Invisible College' [5]. Cross-institution collaboration is increasingly encouraged in British and American universities.

In Britain, traditionally the main sources of research income for universities were the block grants from the University Funding Council (UFC) and Research Council (RC) grants. Since 1988, the system has undergone a radical change. The allocation of funding is made more as a function of performance. Funding is allocated according to criteria which are largely established by a periodic Research Assessment Exercise (RAE). The competitive nature of the public funding allocation process has a number of implications. In particular, a culture of interdisciplinary cooperation is becoming an increasingly significant factor in industrially relevant research and a culture of inter-university collaboration is encouraged.

In the USA, the land-grant university system was founded by Congress during the Civil War to educate working people. Over the last 130 years, the original differences between land-grant universities and other universities have almost disappeared. The growth in the higher education system nationwide and the increasing connectivity across institutions statewide has led to new kinds of land-grant universities. A number of land-grant universities have created branch campuses, and state legislatures have been encouraging cooperation across all higher education institutions within the state.

Proximity to sources of knowledge has been regarded as a source of competitive advantage because of the tactical nature of much know-how, especially for effective technology transfer and knowledge diffusion. Recent evidence shows that the local environment in Scotland, such as its cultural and economic structure, still has a big impact even on the most internationally oriented universities [6]. In this paper, we investigate the use of the WWW in order to achieve a better understanding of the interrelationship between the globalisation of the WWW and the culture of cross-university collaboration in both Scotland and America.

Many studies have demonstrated that there is a strong correlation between citation counts and ratings of academic excellence as measured in terms of research funding, editorships of major journals and peer group ratings [7]. In the UK, the funding bodies for higher education periodically conduct an RAE to evaluate the quality of research in different university departments. Oppenheim [8] showed that there is a correlation between citation counts and the 1992 RAE ratings for British library and information science university departments.

In this paper, we investigate the interaction between the global use of the WWW and the associated cyber-infrastructure for a scientific activity network as described by Latour [4]. In particular, we present two studies, conducted in two different cultural settings, which characterise the profile of a university on the WWW. The first study analyses the connectivity of computer science departmental websites across 13 universities in Scotland, using data automatically gathered from the Internet, and third-party data relevant to research and teaching quality. The second study, independently conducted in the USA, is based on a survey of WWW-based education activities across 50 American land-grant universities. These two studies together provide a richer context of the global use of the WWW in different cultural and economic environments.

2. Study 1: connectivity analysis for Scottish universities

The Scottish universities vary in age, size and prestige. At the time of the study there were 13 universities in Scotland. These institutions obtain the majority of their income from public funds via the Scottish Higher Education Funding Council (SHEFC). SHEFC also funds a number of other smaller specialised colleges. This study was restricted to the SHEFC-funded universities.

2.1. Scottish universities and sample websites

Before 1992, higher education in Scotland was subject to a version of the ‘Binary System’. There were eight Scottish universities, comprising the four ancient universities dating from the middle ages and renaissance, and four others which had their roots in 19th-century colleges, but which only became separate, independent universities in the 1960s. These eight universities received government funding (block grants) for teaching and research through a UK-wide body, the Universities Funding Council (UFC). On the other side of the binary divide there were five ‘Central Institutions’ (CIs), including two polytechnics, which received their block grants from the Scottish Office; these funds covered teaching costs only. Institutions on both sides of the divide could apply for grants to the Research Councils.

In 1992 the Binary System was abolished, the CIs became universities and all universities in Scotland began to receive public funding through a new body, the Scottish Higher Education Funding Council (SHEFC). The new universities (former CIs) could now receive block grants for research. However, the government was moving towards a policy of greater selectivity, which was implemented by means of a RAE conducted by the funding councils, and by a rolling Teaching Quality Assessment (TQA) exercise. In the 1992 RAE, research in each subject in each university was rated on a five-point scale by a peer-review process and the available funds for research were distributed proportionately to the quality of the research and the number of research-active staff. The exercise was repeated in 1996, the scale being extended to seven points and the distribution of funds being even more highly selective. Between 1992 and 1996 teaching in each subject was peer-rated on a four-point, ordinal scale (unsatisfactory, satisfactory, highly satisfactory, excellent).

Internet access for higher education in the UK is provided via the Joint Academic Network operated by UKERNA. There are also four Metropolitan Area Networks for higher education in Scotland, funded by SHEFC. Each of the 13 universities and eight higher education colleges is connected to its local MAN (although in the case of multi-site universities, not every campus is connected). The MANs are known as AbMAN (Aberdeen), FATMAN (Fife and Tayside), EastMAN (East of Scotland) and ClydeNet (the Glasgow conurbation) (Fig. 1).

2.2. Method

Each of the 13 Scottish universities has a computer science or computer studies department. These computer science websites were selected for study because computer science

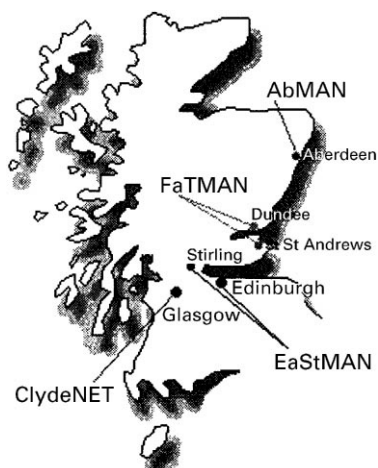


Fig. 1. The Scottish MANs (Source: <http://hel.cee.hw.ac.uk/sman.html>).

departments, in general, have established infrastructure and they are more experienced in developing WWW documents. The connectivity study, called ROADMAP, was conducted between August and September 1996.

2.2.1. Collecting WWW data

In this study, we used the HARVEST system to gather data automatically from the Internet. HARVEST is an integrated set of tools for gathering, extracting, organising, indexing and searching relevant information on the Internet [9]. In particular, two subsystems of HARVEST, Gatherers and Brokers, were used to collect and index subject-specific information on the WWW. We used HARVEST Release 1.4 on an HP workstation with HP-UX 9000 operating system. HARVEST's Gatherers was used to download and digest HTML documents using a modified extraction algorithm, known as a summariser in HARVEST, which determines what types of information should be collected. The boundary of a WWW site was determined by some pattern matching rules which instruct a HARVEST Gatherer to retrieve documents from valid Uniform Resource Locators (URLs) on specific WWW servers.

We have developed a generic approach, called Generalised Similarity Analysis (GSA), for structuring and visualising a hypertext-based information space on the WWW [10]. This approach provides a set of tools for structural analysis and linkage pattern extraction. In this study, HTML documents were automatically retrieved from 13 computer science sites by HARVEST's Gatherers. The traversal concentrated on hypertext reference links. The resultant hypertext reference links, in the form of URLs, were parsed and analysed by hypertext link analysis mechanisms provided by GSA. We analysed patterns associated with incoming and outgoing hyperlinks for each of the 13 computer science sites.

2.2.2. Collecting organisational profile data

Oppenheim [8] showed that there is a correlation between citation counts and the 1992 RAE ratings for British library and information science university departments. In his

study, departments were ranked by citations that each department had received during the same period of time as the RAE92. Departments were also ranked by the average number of citations received by each member of staff in these departments. A Spearman rank order correlation analysis showed that RAE92 ratings strongly correlated with numbers of citations received ($r_s = 0.81, p < 0.01$), and with numbers of citations per member of staff ($r_s = 0.82, p < 0.01$).

Data concerning a hosting department or organisation are called organisational profile variables. These data were gathered from public accessible sources. For example, outcomes of the 1992 and 1996 RAE were published by the Universities Funding Councils [11] and Joint Funding Councils of the United Kingdom [12]. TQA ranks and Student–Staff Ratios (SSR) were obtained from the Times Higher Education Supplement’s Databank [13] on the WWW. Data on research, teaching and equipment funding allocation was obtained from SHEFC’s website. The following data were obtained and used in the study (Table 1).

2.2.3. Connectivity analysis

Data automatically collected from these computer science websites were combined with the available RAE, TQA and funding data for the computer science departments. We traced several types of destinations of hyperlinks associated with a website in order to identify characteristics of a virtual scientific community and profound connections within such a network, including academic-to-academic links, academic-to-commercial links, and international links—notably UK-to-US links—which were further divided into links to academic, commercial, government or military destinations.

For the 13 Scottish sites, we particularly analysed linkage patterns to British commercial sites versus linkage patterns to commercial sites in America. We expected that a two-dimensional framework along the geographical distance and the nature of linked commercial sites may lead to some insights into cultural–technological explanations of the disparity.

We used hypertext link analysis tools developed for GSA to extract all the hypertext links embedded in each HTML document retrieved from a chosen website. All the hypertext links starting from the website were parsed and classified into a number of types by criteria such as protocols used and the nature of target websites.

We then generated a number of matrices and conducted Spearman correlation analysis

Table 1
Organisational profile variables for Scottish universities

| Variable | Definition |
|----------|---|
| RAE92 | Research Assessment Exercise ratings in 1992 for computer science departments |
| RAE96 | Research Assessment Exercise ratings in 1996 for computer science departments |
| ACTIVE | The proportion of active researchers in each department as measured by RAE96 |
| TQA | Teaching quality assessment ranks for each university |
| SSR | Average student–staff ratios across universities |
| RF | SHEFC Research funding level across universities |
| TF | SHEFC Teaching funding level across universities |
| EF | SHEFC Equipment funding level across universities |

for a number of combinations between site connectivity variables and organisational profile variables. The conventional significance level of 0.05 was chosen.

2.3. Results

This section focuses on two types of results of the connectivity analysis. We generated a number of descriptive statistics about the 13 websites and patterns of hypertext reference links associated to these websites. We also derived Spearman correlation coefficients between a number of hypertext linkage patterns and underlying organisational environments.

We focused on the interconnectivity of these websites within the backbone metropolitan networks in Scotland. As an analogue to citation index, we examined the number of incoming links to each website as an indicator of the quality of the information on that site. The number of outgoing links from each website was also considered as an indicator of how well that site disseminated referential information on the WWW.

2.3.1. Overall hypertext references

The average number of HTML documents successfully retrieved from a website in this sample was 1672. The average number of errors which occurred was 501; the most common cause was due to an obsolete URL. Most site-visits by HARVEST Gatherers were scheduled over weekends or between midnight and early morning, local time. The average time taken to collect data from one site was 244.54 minutes, which was about 4 hours. The average number of URLs that Gatherers attempted to resolve was 3653, suggesting that a considerable number of documents were either not in HTML or not accessible.

The first part of a URL identifies the protocol for document transfer between a Web server and a user's computer. The most popular ones are Hypertext Transfer Protocol (HTTP) and File Transfer Protocol (FTP). HTML documents are transferred via HTTP protocol. In our study, 92% of the hyperlinks found in HTML documents at the 13 websites used HTTP. Four percent used FTP and 2% used GOPHER. Other protocols such as NEWS were rarely used.

Spearman correlation analysis showed that the number of hypermedia links (HTTP) provided by a website is positively correlated with the research and teaching assessment scores of the hosting department (see Table 2). On the other hand, the number of

Table 2
Correlation between hypertext references and organisational profile ($p < 0.05$)

| Hypertext references | Organisational profile | Spearman correlation coefficients r_s | Sig. (one-tailed) | <i>N</i> |
|----------------------|------------------------|---|-------------------|----------|
| http links | TQA | 0.819 | 0.000 | 13 |
| | RF | 0.792 | 0.001 | 13 |
| | RAE92 | 0.602 | 0.015 | 13 |
| | RAE96 | 0.563 | 0.028 | 12 |
| | SSR | -0.557 | 0.024 | 13 |
| | TF | -0.792 | 0.001 | 13 |

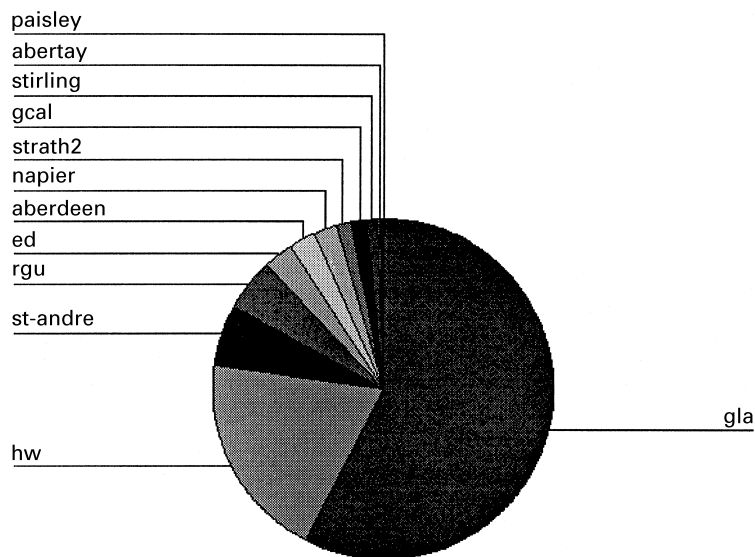


Fig. 2. The number of outgoing hyperlinks from each of the 13 sites.

hypermedia links is negatively correlated with SSRs and the proportion of teaching funding. For example, a lower number of hypermedia links is likely to be associated with a higher SSR.

2.3.2. Outgoing hypertext links

Outgoing links from one of the 13 academic sites to websites outside the university are also called remote hypertext references. Users may find it very useful to have a rich collection of remote hypertext references as a gateway to a wider range of information resources on the WWW. Many people's home pages contain such remote links to further information resources on specific topics. Fig. 2 is a pie chart of the number of remote links found at each of the 13 computer science departmental websites in Scotland. For example, Glasgow University had the largest number of outgoing links to external websites.

We found that the number of remote links is correlated with TQA ranks. It is also positively correlated with SHEFC research and equipment funding levels. A site with a large number of remote links is likely to be associated with a department with a higher

Table 3

Correlation between remote outgoing links and organisational profile ($p < 0.05$)

| Remote hypertext references | Organisational profile | Spearman correlation coefficients r_s | Sig. (one-tailed) | N |
|-----------------------------|------------------------|---|-------------------|-----|
| Outlinks | TQA | 0.683 | 0.005 | 13 |
| | RF | 0.586 | 0.018 | 13 |
| | EF | 0.485 | 0.047 | 13 |
| | TF | -0.608 | 0.014 | 13 |

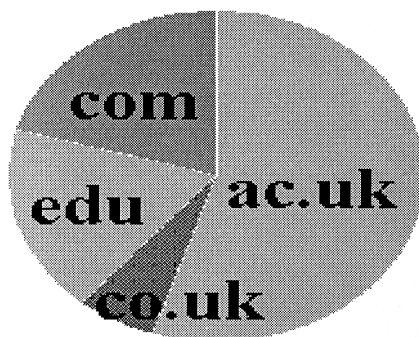


Fig. 3. Distribution of remote outgoing links by destination.

TQA, a large amount of SHEFC research and equipment funding, but a relatively small amount of SHEFC teaching funding, which essentially characterises a strong computer science department in an old university in Scotland (Table 3).

2.3.3. Linkage patterns to academic sites

We found that the majority of remote outgoing links from the 13 sites connected to academic sites had the suffix .ac.uk in their domain names. There were about the same number of remote outgoing links to American academic sites (.edu) and American commercial sites (.com) (see Fig. 3). Remote links to UK companies (.co.uk) were considerably fewer. As for links connecting to academic sites in Britain, about 70% of them were linked to Scottish universities on Scottish Metropolitan Networks (MANs) and the rest were mainly connected to universities in England and Wales.

Spearman correlation analysis revealed a similar pattern regarding the correlation between the number of remote academic links and organisational environment variables. Older universities appeared to exploit the WWW better in establishing academic links (Table 4).

2.3.4. Linkage patterns to American and British commercial sites

In the sector of higher education, both in America and Britain, governments and funding bodies increasingly encourage academic-industry and cross-institution collaboration in order to forge technology transfer and knowledge diffusion effectively.

Table 4

Correlation between references to academic sites and organisational profile ($p < 0.05$)

| Hypertext references to academic sites | Organisational profile | Spearman correlation coefficients r_s | Sig. (one-tailed) | N |
|--|------------------------|---|-------------------|-----|
| ac.uk | TQA | 0.853 | 0.000 | 13 |
| | RF | 0.820 | 0.000 | 13 |
| | RAE92 | 0.619 | 0.012 | 13 |
| | RAE96 | 0.559 | 0.029 | 12 |
| | SSR | -0.546 | 0.027 | 13 |
| | TF | -0.812 | 0.000 | 13 |

Table 5

Top 10 most popular American and British sites (August and September 1996)

| Rank | US site | Count | Type | UK site | Count | Type |
|------|------------------|-------|---------------|----------------|-------|-------------|
| 1 | Java | 188 | Software | Demon | 63 | ISP |
| 2 | Yahoo | 124 | Search engine | Telegraph | 27 | Media |
| 3 | AltaVista | 34 | Search engine | Bookshop | 35 | Publisher |
| 4 | Lycos | 29 | Search engine | Web13/Future | 28/28 | ISP/media |
| 5 | Microsoft | 30 | Software | Cityscape | 26 | Media |
| 6 | AT&T research | 51 | Research | Nexor | 34 | ISP |
| 7 | Netscape | 32 | Software | OUP | 16 | Publisher |
| 8 | NBA | 36 | Sport/music | Almac | 16 | ISP |
| 9 | Digital research | 44 | Research | Musicbase | 18 | Sport/music |
| 10 | Lights | 36 | Others | Virgin Records | 17 | Sport/music |

Table 5 shows the top 10 American and British commercial websites most frequently cited by the 13 Scottish computer science sites. A commercial site was identified by its domain name, i.e. .com for an American site and .co.uk for a British site. These sites were ranked by the number of unique Scottish sites that linked to them. For example, 10 out of 13 Scottish sites had links to *Java Development Toolkit* at java.sun.com at the time of the analysis.

An interesting pattern emerged. Links to American sites were predominated by companies providing leading Internet-related technologies and services, such as *Java programming tools*, *Yahoo* and *AltaVista*. On the other hand, links to British sites were predominated by mass media and entertainment such as *Daily Telegraph* and *Channel 4* (at www.cityscape.co.uk).

We generated a profile of the top-ranked popular commercial sites in a two-dimensional configuration using Multidimensional Scaling (MDS). Each of these sites was represented by a vector according to how frequently it was referenced across the 13 Scottish sites. The frequencies were standardised over all the Scottish sites so as to minimise the bias towards large sites in Scotland. Fig. 4 shows the resultant configuration generated by SPSS. This spatial configuration accounts for 85% of the variance. We annotated the configuration with some lines to highlight sites that are similar to each other.

There are usually no hard and fast rules to interpret the nature of each dimension appearing in a MDS configuration. Nevertheless, by studying a given spatial layout, one can identify significant patterns that may reflect effects of some underlying factors. In Fig. 4, for example, Dimension 1 highlighted the contrast between scientific work at one end with research laboratories in large American companies and two sites particularly devoted to music at the other. Dimension 2 may indicate the generality or transferability of the content of a site with respect to particular cultures. In fact, three notable sites were placed along this dimension: *Java programming*, *generic search* and *Daily Telegraph*. The positions of *Yahoo* and *AltaVista* suggested some connections to the generic nature of their indexing and search facilities. On the other hand, popular British commercial sites were strongly characterised by the British culture, for example, *Daily Telegraph* (one of the most popular quality newspapers in Britain), *Channel 4* (a nationwide TV channel) and *Oxford University Press* (OUP).

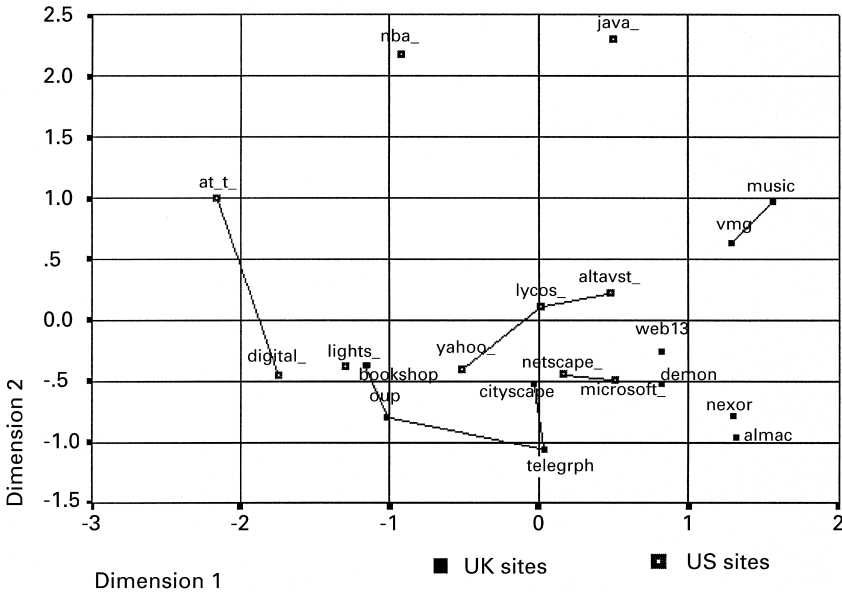


Fig. 4. The profile of the most popular American and British commercial sites to Scottish universities (MDS stress = 0.184, RSQ = 0.846).

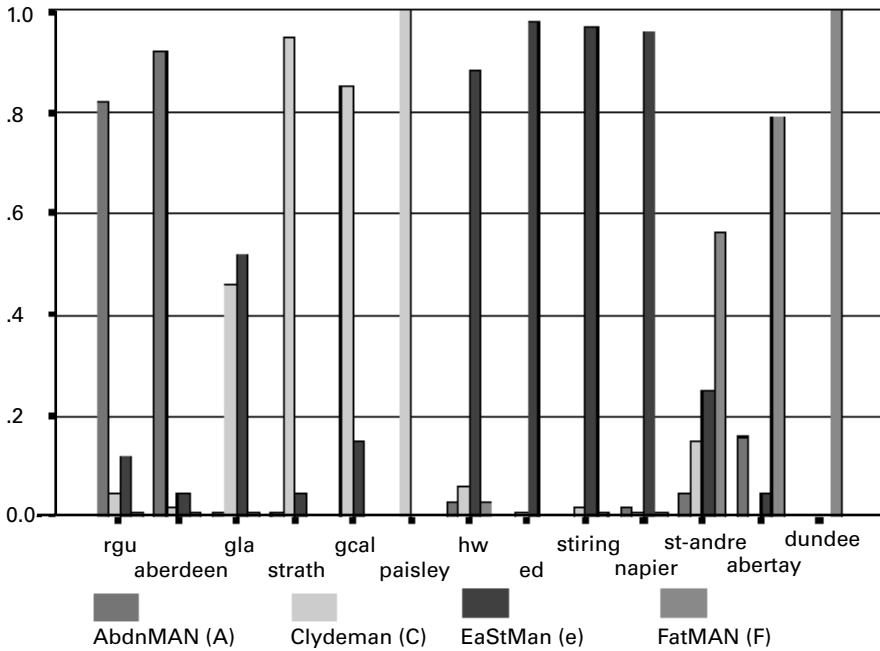


Fig. 5. Cross metropolitan networks links at each of the 13 Scottish sites. Some sites may not have links to all the metropolitan networks.

Table 6

Correlation between reference locality and organisational profile ($p < 0.05$)

| Hypertext reference locality | Organisational profile | Spearman correlation coefficients r_s | Sig. (one-tailed) | N |
|------------------------------|------------------------|---|-------------------|-----|
| mans/ac.uk ratios | RAE96 | -0.723 | 0.004 | 12 |
| | SSR | 0.528 | 0.032 | 13 |

2.3.5. Hypertext reference locality

Hypertext reference locality is the ratio of the number of outgoing links to Scottish universities to the total number of outgoing links to British universities as a whole. A site with a high locality ratio may have concentrated more on activities in Scottish universities than elsewhere. Indeed, many outgoing links connected universities on the same metropolitan network. On the other hand, at some sites we found a considerable amount of cross-metropolitan network links. For example, Glasgow University on a metropolitan network called ClydeNet (C in Fig. 5), had a substantial number of links to Edinburgh University, which is on a metropolitan network known as EaStMAN (E in Fig. 5).

We found that the locality ratio strongly correlated with the latest RAE ratings in 1996 ($r_s = -0.723, p = 0.004$) (Table 6). A department with a higher rating in RAE96 tends to have connected to a wider range of universities outside Scotland. In general, a high RAE96 rating and a low SSR typically characterise a strong research department in an old university.

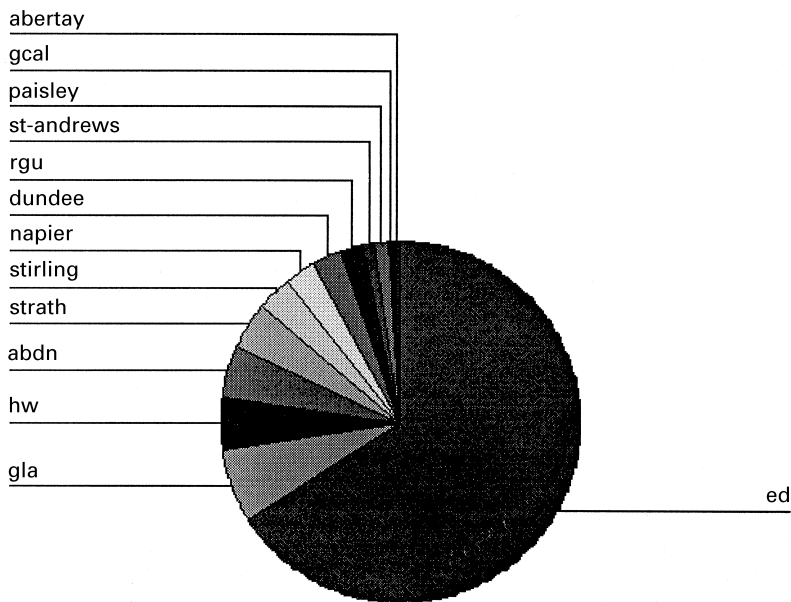


Fig. 6. Hypertext citations received by each of the 13 sites from other computer science sites in Scotland.

Table 7

Correlation between hypertext citations per site and organisational profile ($p < 0.05$)

| Hypertext citations received per site | Organisational profile | Spearman correlation coefficients r_s | Sig. (one-tailed) | N |
|---------------------------------------|------------------------|---|-------------------|----|
| inlinks | RF | 0.751 | 0.002 | 13 |
| | TQA | 0.714 | 0.003 | 13 |
| | RAE92 | 0.602 | 0.015 | 13 |
| | SSR | -0.537 | 0.029 | 13 |
| | TF | -0.762 | 0.001 | 13 |

2.3.6. Incoming links to Scottish sites

In this study, the number of incoming links to a site was also called hypertext citations. Hypertext citations, by analogue to bibliographic citations in Library and Information Science, indicate the impact of a website on peer researchers and teaching staff in other Scottish universities. By the same principles as described by Latour [4], a department should aim to achieve a higher hypertext citation count as they would with scientific publications in traditional media (Fig. 6).

Spearman correlation analysis indicated that hypertext citations strongly correlated with SHEFC research funding levels, SHEFC teaching quality assessment and RAE ratings in

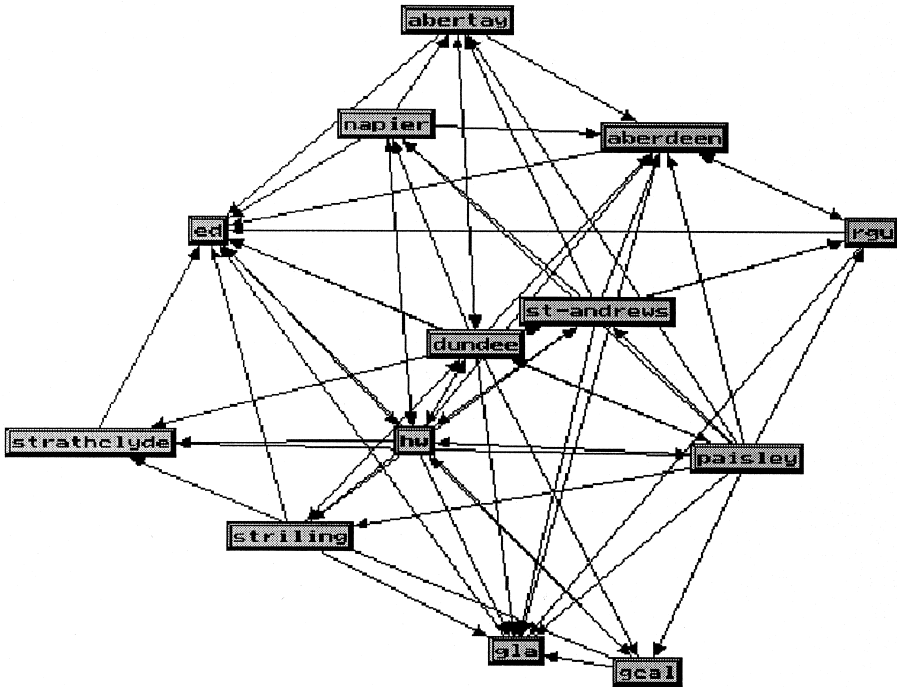


Fig. 7. A connectivity map of the 13 university departmental websites. (Adapted from Ref. [10].)

Table 8

Correlation between inter-site connectivity and organisational profile ($p < 0.05$)

| Intersite connectivity | Organisational profile | Spearman correlation coefficients r_s | Sig. (one-tailed) | N |
|------------------------|------------------------|---|-------------------|-----|
| ln(in*out) | TQA | 0.664 | 0.007 | 13 |
| | RF | 0.611 | 0.013 | 13 |
| | TF | -0.622 | 0.012 | 13 |

1992. Sites associated with a high SSR and a high proportion of SHEFC teaching funding tended to receive fewer hypertext citations (Table 7).

2.3.7. Inter-site connectivity

We used hypertext linkage analysis tools developed for Generalised Similarity Analysis [10] and generated a connectivity map of all the 13 university departmental websites (see Fig. 7). This connectivity map was generated so that if two sites are highly connected to each other, then they will be placed close to each other on the map. The direction of a link on the map indicates the collective direction of hypertext reference links (Fig. 7).

We defined inter-site connectivity as $\ln(\text{incoming_links} \times \text{outgoing_links})$ so as to produce a balanced view of hypertext citations and hypertext references associated with each site. The intersite connectivity is positively correlated with SHEFC teaching quality assessment and research funding and negatively correlated with SHEFC teaching funding. In essence, this result is consistent with Spearman correlation analysis on other attributes of a website discussed in previous sections (Table 8).

We ranked the 13 computer science websites in Scotland by corresponding intersite connectivity (see Table 9). The ranking was based on the data collected between August and September 1996. Since the rapid growth of the WWW, these websites are likely to

Table 9

Computer science departmental websites ranked by intersite connectivity

| Site ID | University computer science depts. | Intersite connectivity ^a $\ln(\text{incoming} \times \text{outgoing})$ | Rank |
|---------|------------------------------------|--|------|
| 9 | Heriot-Watt | 12.95 | 1 |
| 8 | Glasgow | 11.02 | 2 |
| 5 | Edinburgh | 10.70 | 3 |
| 15 | Robert Gordon | 9.91 | 4 |
| 21 | Strathclyde | 9.67 | 5 |
| 1 | Aberdeen | 8.93 | 6 |
| 20 | Stirling | 8.32 | 7 |
| 11 | Napier | 8.27 | 8 |
| 18 | St Andrews | 7.37 | 9 |
| 3 | Dundee | 7.13 | 10 |
| 6 | Glasgow Caledonian | 5.72 | 11 |
| 2 | Abertay Dundee | 5.54 | 12 |
| 13 | Paisley | 0.00 | 13 |

^aAs in August–September 1996.

have changed dramatically over the last few months. In the future, it would be an interesting topic to repeat the ranking periodically and monitor systematic changes in the patterns of connectivity.

3. Study 2: content analysis for American land-grant universities

The WWW enables users to access geographically distributed information resource. Using the WWW for education has been widely exploited in different countries. The connectivity analysis introduced in this paper has the advantage of being systematic and completely repeatable periodically at little cost. On the other hand, there are many subtle and perhaps important features of various universities, especially across different countries. We contrast the connectivity analysis in Scotland with a content analysis of WWW-based education activities in American universities to establish a balanced context for understanding cultural–technological implications of the globalisation.

This content analysis was independently conducted on WWW educational activity at American land-grant universities. Some preliminary results of its pilot study were reported in Ref. [14]. A number of relationships between institutional attributes and the amount of WWW activity were investigated.

3.1. Land-grant universities

The American land-grant university system was founded by Congress during the Civil War to educate working people. Over the last 130 years, the original differences between land grants and other universities have almost disappeared. In our second study, a land-grant university is defined as any university in the appropriately titled list from the US Department of Agriculture [15].

The growth in the higher education system nationwide and the increasing connectivity across institutions statewide has led to new kinds of land-grant universities. A number of land-grant universities have created branch campuses, and state legislatures have been encouraging cooperation across all higher education institutions within the state. Some statewide university systems now include the traditional land-grant university and other higher education institutions, such as community colleges.

The WWW offers a new and rich source of information for increasing access, improving quality, and constraining per student costs in universities. Which universities have the most reason to want to increase access and might such universities also be the ones that use the WWW most? Johnson et al. [16] reviewed the WWW home pages of many universities and determined for each university who created the home page and the types of information available.

3.2. Method

Student analysts were recruited at Washington State University for casual work at \$7.50 per hour to carry out a survey on WWW-based course materials. They worked in the office of Institutional Research under the supervision of permanent staff.

From the US Department of Agriculture [15] list of land-grant colleges and universities in the USA, 50 universities were selected. This list orders institutions within a state so that in effect the largest institution of the state is listed first. For this study, the first institution of each state in the list was chosen.

The survey of WWW-based course materials at a land-grant university started from the university's home page and a typical path was recommended to analysts to follow from academics, to colleges, to departments, and finally to courses. On the other hand, analysts were encouraged to use whatever traversals they thought would lead to the most course-specific information.

Each analyst was assigned to a multimedia, networked computer and given instructions on how to record observations. Analysts collected data into an Excel spreadsheet that had been structured for this purpose. Each analyst was also given a copy of the Classification of Instructional Program (CIP) codes [17] of the US Department of Education and an electronic pointer to the home page of the university to be surveyed. For each visited file with educational content the analyst recorded, in fixed fields of the spreadsheet, the WWW address of the file, the score on each of five attributes, and the CIP code based on the presence or absence of educational materials.

On finding course-specific information, the analysts assessed five attributes of WWW content of that course. The attributes were for the presence of syllabi, course material, hypermedia links, online interaction, and online submissions:

1. Syllabus: the presence of syllabus, course outline, course overview, seminar schedule.
2. Course materials: the presence of assigned readings, study guides, study questions and answers, electronic handouts, etc. The idea is that course-related information be available on-line. A reference list of assigned or suggested readings does not constitute course material.
3. Hypermedia links: links to supporting data are embedded in the HTML source code.
4. On-line interaction: electronic classrooms, bulletin boards, or discussion forums dedicated to that course specifically. Simple e-mail does not count.
5. On-line submission: the ability to submit homework, quizzes, papers, exams, etc. on-line.

For several universities, more than one analyst assessed the university. If the results for the same university were significantly divergent for two different analysts, then the senior staff examined the differences and perfected the results.

A university's formal technology policies, provision of hardware and software, faculty development opportunities, and technical and administrative support have been claimed by others to influence a university's level of educational technology activity. Many attribute values for each university, such as total enrolment, instructional support dollars, and academic support dollars are available from the Integrated Post-secondary Education Data System [18]. Although we intended to use the latest data in the analysis, more recent NCES data were not readily available at the time of the analysis. The content analysis tested a correlation relationship between the absolute size of a university and the absolute number of courses it taught with the support of the WWW. Similar tests were conducted on instructional support dollars and academic support dollars.

One might expect that geographically isolated universities would focus more on the WWW. For each home campus of a university in the land-grant university survey, the miles as the crow flies between the home campus and the closest metropolitan area in the state of that university were measured. A regression between these distances and the overall ranking did not reveal a positive relationship.

One might expect that a multi-campus university would find more benefit from WWW courses than a single-campus university. A course that was otherwise only taught on one campus might be taught across all campuses when the WWW is used. Accordingly, the survey was extended to include another 0–1 value: if the university has multiple campuses, it scored a 1 and otherwise a 0.

Each university was assigned a 0 or 1 score based on the following criteria:

- if the university had two or more campuses, then it earned a score of 1, but
- if the university had only one campus, then it earned a score of 0.

The analysis tested whether the multi-campus score and the overall ratings of WWW-based education activity are correlated.

There are some apparent connections between these attributes in American universities and their counterparts in Scottish universities in Study 1. For example, SSRs and funding levels correspond to university size and instructional support dollars and academic support dollars, whereas multi-campus scores may correspond to cross-MAN linkage patterns in the study of Scottish universities.

3.3. Results

The universities can be ranked according to scores on each attribute. For an overall ranking, a weighting scheme was devised. Attributes 4 and 5 form group C, attributes 2 and 3 form group B, and attribute 1 is alone in group A (Table 10). Group C, for example, is weighted by a factor of 3 in the overall ranking because it represents a more significant use of the WWW for education than does group B or A. The final value for a university is a weighted sum of these three groups. Data for the top 15 universities by the overall ranking are given in Fig. 4.

When the total enrolment data from IPEDS was regressed with the overall ranking, no significant relationship was established. Similarly, the data on instructional support dollars and academic support dollars did not correlate with the overall ranking. No significant positive or negative relation was found between the percentage of tenured faculty and its overall rank.

The top half of the overall ranking scores was assigned a 1 and the bottom half a 0 in order to test a simple correlation between the ranking scores and the 'multi-campus score'. Using the Microsoft Excel CORREL function, the correlation was -0.06 and indicates no predictive relationship between the two attributes.

The following six programs were highly ranked across the 50 universities, which accounted for 4219 (or 60%) of the grand total sum of attribute scores:

1. Computer and Information Engineering;
2. Engineering;
3. Life Sciences;

Table 10
The top 15 universities for their significant WWW-based education activities

| University | Syll. (A) | Rank | Mat. | Link | Total (B) | Rank | Int. | Sub. | Total (C) | Rank | 1A + 2B + 3C | Rank |
|--------------|--------------|------|------|------|--------------|------|------|------|--------------|------|--------------------|------|
| Pennsylvania | 215 | 2 | 91 | 60 | 151 | 1 | 13 | 16 | 29 | 3 | 604 | 1 |
| Virginia | 186 | 3 | 103 | 19 | 122 | 5 | 2 | 8 | 10 | 15 | 460 | 2 |
| New Jersey | 138 | 7 | 102 | 37 | 139 | 3 | 3 | 2 | 5 | 22 | 431 | 3 |
| Missouri | 97 | 12 | 86 | 62 | 148 | 2 | 5 | 6 | 11 | 12 | 426 | 4 |
| Maryland | 317 | 1 | 46 | 4 | 50 | 14 | 1 | 1 | 2 | 27 | 423 | 5 |
| New York | 116 | 9 | 88 | 42 | 130 | 4 | 8 | 1 | 9 | 17 | 403 | 6 |
| Kansas | 107 | 10 | 108 | 5 | 113 | 6 | 9 | 11 | 20 | 5 | 393 | 7 |
| Washington | 119 | 8 | 66 | 45 | 111 | 7 | 8 | 5 | 13 | 9 | 380 | 8 |
| Florida | 80 | 15 | 78 | 8 | 86 | 9 | 18 | 16 | 34 | 2 | 354 | 9 |
| Colorado | 105 | 11 | 45 | 20 | 65 | 12 | 36 | 3 | 39 | 1 | 352 | 10 |
| California | 171 | 4 | 73 | 7 | 80 | 11 | 2 | 0 | 2 | 28 | 337 | 11 |
| Connecticut | 58 | 18 | 65 | 28 | 93 | 8 | 4 | 7 | 11 | 11 | 277 | 12 |
| Oregon | 57 | 20 | 53 | 28 | 81 | 10 | 0 | 0 | 0 | 47 | 219 | 13 |
| Minnesota | 65 | 16 | 41 | 16 | 57 | 13 | 4 | 6 | 10 | 13 | 209 | 14 |
| Utah | 149 | 5 | 8 | 8 | 16 | 31 | 0 | 1 | 1 | 33 | 184 | 15 |

Syll.: syllabus, Mat.: content material, Link: hyperlink, Int.: interaction, Sub.: submission.

4. Physical Sciences;
5. Business Management;
6. Agricultural Sciences.

WWW courses in American land-grant universities were predominated by technologically-sophisticated programs, such as computer science and engineering. These programs are often the most active in exploiting the WWW in teaching and learning.

4. Discussion

The Scottish study indicated several strong correlation relationships between structural connectivity measures of a computer science website and the longer-term organisational profile of the department and its hosting institution. For example, it was shown that, similar to Oppenheim's (1995) study, there is a strong, positive Spearman correlation between the number of incoming links attracted by a site and the RAE ratings of its hosting department. We also found that the intersite connectivity of a website was strongly correlated with teaching quality assessment and research funding associated with the corresponding university department.

The historical impact of the Scottish university system is reflected in these patterns. Old universities, which generally have lower SSRs and higher research funding levels than new universities, tend to be more active in the use of the WWW with a higher intersite connectivity score.

Scottish universities, in general, established strong connections to American commercial sites such as *Yahoo* and *AltaVista*, and sites with advanced research in large American

companies' research laboratories. The fact that java.sun.com was one of the most popular sites among Scottish computer science sites underlined the widespread interest in Java programming. On the other hand, links to British commercial sites were essentially shaped by *Daily Telegraph*, *Channel 4* and *Oxford University Press*. These patterns highlighted the impact of global technology transfer and knowledge diffusion across different cultures. To further understand the nature of WWW-based activities, one may need to investigate interrelationships among users' individual differences, interactive tasks and appropriate enabling technologies in the use of the global and distributed hypertext [19].

There are some apparent anomalies in the Scottish findings: for example, the University of Edinburgh, where the Computing Science Department was highly rated in the 1992 and 1996 RAE, as well as in the TQA, has attracted the greatest number of incoming hypertext links from other Scottish universities (incoming links = 885). But surprisingly, it made much fewer hypertext references to the rest of the websites in Scotland (outgoing links = 39), which was the fifth largest website in terms of the number of outgoing links. Indeed it had fewer incoming links than some former CIs. We found that, having already established itself as a world class department in computing, it has directed its efforts to strengthen its connectivity with other world class universities such as Carnegie Mellon University and MIT. This may suggest that different perceptions of the appropriateness of technology and user acceptance of the WWW are ongoing in the different departments. To maximise the potential benefits from this new information medium, universities must deliberately mobilise and dedicate high quality human and material resources in aid of the 'outside scientists'.

Study 2, the American land-grant university study, does not show such strong connections between the size of a university and patterns in their educational activities on the WWW. One explanation could be that American land-grant universities are in general more homogeneous to each other than their counterparts in Scotland, the WWW use might not lead to a contrast as sharp as between old and new universities in Scotland. Organisational profile attributes in our investigation are subject to cultural and micro-political influences. For example, research and teaching funding levels reflect differential material resources historically available on either side of the binary divide. In addition, they directly reflect government and funding bodies' policies in a larger cultural and economic environment. In the future, we would be particularly interested in conducting similar studies on computer science websites in England as their departments are affected by more similar organisational factors such as RAE and old–new university establishments. Further studies are needed for a better understanding of cultural, political and economical impacts on deeper globalisation of the enabling techniques on the Internet, for example, by examining other dimensions of cultural bases identified in Day [1].

5. Conclusion

Various patterns in hypertext references and hypertext citations indicate underlying relationships between the use of the WWW and some characteristics reflected in the profile of an organisation. That is the old university department, which received more research funding and smaller proportions of teaching funding, good RAE ratings and SHEFC

teaching assessment ranks, is more likely to have a balanced and well-developed website than a new university department, which can be characterised towards the other end of the scale for almost all the organisational environment variables (especially with larger SSRs and less research funds from SHEFC).

To exploit the WWW in innovative teaching and learning as the mission of universities in Scotland, one should provide additional help to new universities in taking up the globalised technology. In fact, several SHEFC funded projects aim to increase the awareness of Scottish MANs and using the MANs in cross-institution teaching with technologies such as video-conferencing and Web-based courses. A number of centralised WWW-based learning centres such as Clyde Virtual University have proven to be very useful vehicles for lecturers and students who are short of resources and technical supports. The success of such initiatives may well depend on complex interactions between the globalisation of enabling information and telecommunication technologies and long-term, profound cultural and economic environments.

Acknowledgements

The authors wish to thank Professor Terry Mayes at Glasgow Caledonian University for discussions on Internet resource discovery projects in Scotland. Also to acknowledge the contributions to the survey of the USA land-grant universities of Jim Rimpau, Jerry Gordon, Todd Sansom, Mark Wagner, Curtis Bowman, and Tom George and the generous support of Washington State University.

References

- [1] D. Day, Cultural bases of interface acceptance: Foundations, in: Proc. of HCI '96, London, UK, August, Springer-Verlag, London, 1996, pp. 35–47.
- [2] J.E. Pitkow, C.M. Kehoe, Emerging trends in the WWW user population, *Communications of the ACM* 39 (6) (1994) 106–108.
- [3] C. Demello, University and college home pages: survey results, <http://www.mit.edu:8001/people/cdemello/results.html>, 1996.
- [4] B. Latour, *Science in Action: How to Follow Scientists and Engineers Through Society*, Open University Press, Milton Keynes, UK, 1987.
- [5] D. Crane, *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*, Chicago University Press, Chicago and London, 1972.
- [6] Scottish Enterprise, Commercialisation enquiry: final research report, Scottish Enterprise and Royal Society of Edinburgh, 1996.
- [7] E. Garfield, *Citation Indexing: Its Theory and Application in Science, Technology, and Humanities*, Wiley Interscience, New York, 1979.
- [8] C. Oppenheim, The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British library and information science university departments, *Journal of Documentation* 51 (1) (1995) 18–27.
- [9] C.M. Bowman, P.B. Danzig, U. Manber, F. Schwartz, Scalable Internet resource discovery: research problems and approaches, *Communications of the ACM* 37 (8) (1994) 98–107.
- [10] C. Chen, Structuring and visualising the WWW by generalised similarity analysis, in: Proc. Hypertext '97, Southampton, UK, 6–11 April, ACM Press, 1997, pp. 177–186.

- [11] Universities Funding Council (UFC), Research Assessment Exercise 1992: The Outcome, Universities Funding Council, Circular 26/92, Bristol, 1992.
- [12] Joint Funding Councils of the United Kingdom (JFCUK), 1996 Research Assessment Exercise: the outcome, <http://back.niss.ac.uk/education/hefc/rae96/>, 1997.
- [13] Times Higher Education Supplement (THES), Teaching quality assessments, <http://www.timeshigher.newsint.co.uk/DATABANK/databank.html>, 1996.
- [14] R. Rada, J. Rimpau, C. Bowman, J. Gordon, T. Henderson, T. Sansom, WWW activity and the virtual university, *Educational Technology and Management* 36 (5) (1996) 49–51.
- [15] US Department of Agriculture (USDA), Land Grant Colleges and Universities, US Department of Agriculture, 1995.
- [16] A. Johnson, W. Mao, S. Mishler, V. Popov, A survey of college and university WWW sites, http://www.ilt.columbia.edu/academic/classes/TU5020/projects/he/higher_ed.html, 1996.
- [17] R. Morgan, E. Hunt, J.M. Carpenter, Classification of Instructional Programs (1990 ed.), US Department of Education, National Centre for Education Statistics (NCES), 1990.
- [18] National Centre for Educational Statistics (NCES), Integrated Postsecondary Education Data System 1992, CD-ROM, US Department of Education, Office of Educational Research and Improvement, 1992.
- [19] C. Chen, R. Rada, Interacting with hypertext: a meta-analysis of experimental studies, *Human-Computer Interaction* 6 (2) (1996) 125–156.