Which English Dominates the World Wide Web, British or American?

Eric Atwell,[1] Junaid Arshad,[1] Chien-Ming Lai,[1] Lan Nim,[1]
Noushin Rezapour Asheghi,[1] Josiah Wang[1] and Justin Washtell[1]

1. Introduction

This paper reports the results of an experiment to combine research and teaching in Corpus Linguistics, using an AI-inspired intelligent agent architecture, but casting students as the intelligent agents (Atwell 2007). Computing students studying Computational Modelling and Technologies for Knowledge Management were given the data-mining coursework task of harvesting and analysing a Data Warehouse from WWW, using WWW-BootCat web-as-corpus technology (Baroni et al 2006). Each student/agent collected English-language web-pages from a specific national top-level domain, and the analysis task involved comparing their national web-as-corpus with given "gold standard" samples from UK and US domains, to assess whether national WWW English terminology/ontology was closer to UK or US English. Results from 93 countries worldwide were collated to give an overview answer to the question: Which English dominates the World Wide Web, British or American?

2. Methods

The detailed coursework specification is given in Appendix A. The task was cast as an exercise in applying the CRISP-DM methodology for computational modelling: the Cross-Industry Standard Process for Data Mining projects. The CRISP-DM methodology specifies a series of phases or sub-tasks in a data-mining project; it is a "recipe" to follow, allowing novices and non-experts to carry out data mining experiments successfully. The students' success in carrying out the exercise is a testament to the practical value of the CRISP-DM methodology.

The World Wide Web is divided into national domains, which makes it easy to collect a corpus of English-language web-pages from a specific country: Google has Advanced Search options to restrict results to a specified domain and language; WWW-BootCat uses Google to search for web-pages, and allows users to specify these options. English is in effect a minority language on the WWW, in that a majority of web-pages world-wide are in languages other than English; however, most national domains do include a large amount of English, showing that English is a truly international language. Our survey was not restricted to countries where English is a native language. We tried to include a wide variety of countries, and we succeeded in collecting 200,000-word samples from most national domain. The exceptions were either very small national domains (e.g. South Georgia Island), or countries with legislation favouring a language other than English (e.g. Algeria has laws promoting publication in Arabic over ex-colonial French, and as a side-effect these have also curtailed the use of English).

[1] School of Computing, University of Leeds
   e-mail: nlp@comp.leeds.ac.uk

Having collected their national English sub-corpus, each agent (student) had to decide whether it was closer to British or American English. Corpus Linguists on the CORPORA email discussion list forewarned us that the task would not be straightforward: many examples of "American" spellings are found in the British National Corpus, so we might have problems with the man-in-the-street assumption that these are two distinct varieties of English. As the students were substituting for AI intelligent agents (and as they were Computing rather than Linguistics students), they could not apply sophisticated linguistic knowledge to the problem. Instead, each student/agent used simple computable measures to compare their national web-as-corpus with given "gold standard" samples from UK and US domains.

The comparison methods included examining Log-Likelihood profiles and averages comparing word-frequencies in domain ,UK and US corpora; counting occurrences of selected words known to have different UK/US spellings (eg color/colour); counting occurrences of concepts realised by different UK/US words (eg fawcet/tap). Analysis was only at the lexical level: we had no means of comparing syntax or looking for characteristically UK v US grammar.

This exercise produced a detailed country-by-country analysis of the results from nearly a hundred student/agents, a large collection of national reports documenting the relative dominance of the two main varieties of English across the World Wide Web. However, although this exercise produced a large volume of "results", it was still difficult to see patterns emerging. As a follow-up exercise, Masters students on the Computational Modelling class were asked to collate and compare results across a group of countries in a single geographical or political region, to produce overviews of English in the region. Students could base their regional overview on the results gathered in the first exercise, though some chose to collate and analyse their own web-as-corpus data afresh. Each regional report was to be written as a research journal paper, targeted at a journal specific to the region. Appendix B shows the detailed specification for this follow-up exercise.

3. Results

The following are summaries of papers written by Masters students:

3.1 Analysis of English used in a web corpus from the Middle East (Junaid Arshad)

The web is huge, free and easily accessible to everyone. It contains billions of words of text and can easily be used for the study of linguistics. Language scientists are becoming increasingly interested in using this rich data source. This study involves the use of the rich data source of the World Wide Web as the core database for our analysis. In this study we present the collection and analysis of data from a subset of World Wide Web i.e. sample WWW texts from the country level domains of the Middle East. The aim of this study is to show that web can be used as a data source for the study of linguistic and analyse the English used in the websites of the Middle East and decide whether the English used in these websites is closer to British English or American English. To achieve this objective the popular data mining approach of CRISP-DM has been used. We collected English web-as-corpus samples from nine Middle East countries, and compared each with Gold Standard English web-as-corpus

samples from UK and US domains, using various comparison methods such as Log-Likelihood frequency profiles. We found that Jordan and Egypt English corpora were closer to UK than US English; English websites in Saudi Arabia, Lebanon, Israel, Kuwait, and Bahrain were more similar to US English than UK English; and UEA and Iran English websites contained a mix of UK and US English, with neither dominant. Our study has resulted in the conclusion that the English used in the web resources of the Middle East is more closely related to American English than British English.

3.2 Studying Influences of British English and American English on World Wide Web in Southeast Asia by Applying Web as Corpus (Chien-Ming Lai)

In this paper, a study of the English texts used in countries in Southeast Asia is presented. The objective is to decide whether the English words used in the chosen country in the region is closer to British English or American English. Computational techniques are applied to collect the materials used in the study from the World Wide Web (WWW). The countries studied were Indonesia, Malaysia, Philippines, Singapore, Thailand and Vietnam. Among these countries, only Philippines and Singapore recognize English as official language, but English is widely used in the other countries. Three analysis works that are similarity comparison, nation-specific terms study and Log-Likelihood comparison are performed to investigate the differences. From the results of the analysis works, the English texts used in most of the chosen countries in the Southeast Asia are closer to the American English. The nation-specific English terms in each country and patterns of the terms are identified, and lists of the specific terms are presented with further discussion as well.

3.3 The Dominant English Type within the World Wide Web Domains of France and its Former Colonies (Lan Nim)

The stark contrast between British and American English can be seen in many books, journals or on the internet. Most countries have their own domain on the World Wide Web (WWW), making it possible to examine which variety of English is preferred in a given national WWW domain, even in countries with a national language other than English. This paper investigates whether the English used in the WWW domains of France (.fr) and its former colonies of Vietnam (.vn), Laos (.ln), Mauritius (.mu) and Senegal (.sn) is dominated by British English, American English or neither?

The decision was taken to make use of the CRISP-DM methodology and Bootcat Corpus Tool because of its ease of use to build a corpus and its ability to process the data quickly. The results were analyzed and formulas were deployed to calculate that British English is more dominant overall in Francophone domains compared to American English. However, some local variation was observed: American English is more widespread in Vietnam, probably due to American political influence after the end of French colonization; and, more surprisingly, American English seems more prevalent than British English in the .FR domain of France.

3.4 Which English dominates the World Wide Web in countries where English is a native language: British or American? (Noushin Rezapour Asheghi)

The aim of this paper is to discover whether the English used on web-sites in countries where English is native language is closer to British English or American English. According to Melchers and Shaw classification of countries by domain of English use, in Australia, New Zealand, Canada, Ireland, South Africa and the Caribbean all the public and private interaction takes place in English for the majority of the population. Using web as corpus approach has increasingly been used to collect text dataset in recent years. In this paper, texts in URLs in domains of each of these countries which represent their English have been the data-source to build a corpus. The idea is to provide evidence showing whether each corpus is closer to British English or American English. The CRISP-DM methodology has been used to derive knowledge from these datasets and then the results of the analysis are presented.

The techniques which have been used in this paper are Log-Likelihood value and the frequency comparison of words spelt differently in British and American English. The results from first technique in modelling phase indicate that English used in Australian, South African and Irish web sites are closer to British English and text in New Zealand, Jamaican and Canadian web sites are more similar to American English. However, there is not a great difference between the results of comparing these corpora with British and American English.

The results of the second method do not completely match the results of the first technique in New Zealand domain. Although, the first approach of comparing New Zealand corpus with British and American corpora shows that this corpus is closer to American English, the result of second method indicates that British spelling is used predominantly in New Zealand domain.

3.5 Dominance of British and American English on the World Wide Web in Malaysia, Singapore and Brunei (Josiah Wang)

This paper discusses the rivalry between British and American English on the World Wide Web, primarily in three countries: Malaysia, Singapore and Brunei. We first provide a brief overview of these countries in terms of their geographical placement, their history as British post-colonial countries, and their usage of English. We then proceed to examine the dominance of British or American English on the Internet in these countries. The Web is used as a corpus to determine whether English text on Malaysian, Singaporean and Bruneian web-sites are closer to British or American English. As a comparison, we have also included three other neighbouring countries in this study: Indonesia, Papua New Guinea and the Philippines. Random datasets are collected to represent English texts from each of these countries. These texts are compared with English corpora from U.K. and U.S. web-pages. The corpora are compared using three methods: (1) determining overlapping word frequencies, (2) determining common words, and (3) examining significant words with log-likelihood statistics. The results from this comparison are used to gauge whether the influence of British or American English in these countries extends to the Internet.

Our results are consistent with our hypothesis; Former British colonies like Malaysia, Singapore and Brunei still favour British English on the World Wide Web. In addition, Indonesia and Papua New Guinea which are indirectly influenced by British English (i.e. through the Netherlands and Australia) also tend to lean towards British English. The Philippines on the other hand still continue to exhibit America's influence with their preference to American English on the Internet.

3.6 The Polynesian influence on English in the World Wide Web of Pacific island nations (Justin Washtell)

This study analyses the effect of indigenous Polynesian languages upon the balance of a core of function (non-lexical) words in sample English web corpora taken from Polynesian island nation domains. Sample corpora are captured from a selection of New Zealand, Cook Islands and French Polynesian websites. These corpora are compared to those recovered from .uk and .us domains and significant grammatical differences are identified. Noted differences are compared with those found between a French corpus from France and one captured from French Polynesian websites using an identical technique. The findings are used to assess the theory that identified influences stem from a Polynesian influence upon European languages use, as opposed to being manifestations of some arbitrary large-scale geographic variation in English usage.

4. Conclusions

At the outset, we expected American English to dominate the WWW: computing generally has been American-led; and multinational companies with national branches might be expected to base their English-language pages on American "originals". We were pleasantly surprised to find that UK English is holding its own on the WWW, and even preferred over US English in many domains and most larger regions except for North and South America.

However, we also had an unforeseen finding: that often it was difficult to see any clear preference for British or American English, at least on the basis of the straightforward computational metrics available. Although intuitively there does seem to be a clear difference between the two varieties, in practice this actually affects only a very small proportion of words in web-pages. The most noticeable difference between British and American English is in pronunciation, which of course is not apparent in web-pages.

We hope to run a similar exercise for next year's classes. However, we need to find an alternative research question to investigate, with clearer metrics, which we can hope to answer more clearly. It seems that neither British nor American English is overwhelmingly dominant on the World Wide Web.

References

Atwell, Eric. Combinatory Hybrid Elementary Analysis of Text. Submitted to Corpus Linguistics 2007.

Baroni, Marco; Kilgarriff, Adam; Pomikalek, Jan; Rychly, Pavel. *WebBootCaT: instant domain-specific corpora to support human translators.* In Proceedings of EAMT 2006 - 11th Annual Conference of the European Association for Machine Translation. 2006.

Appendix A: Computational Modelling first coursework assignment
(lecturer: Eric Atwell, School of Computing, University of Leeds)

Which English dominates the World Wide Web: British or American?

This assignment involves the collection and analysis of data from your chosen domain or subset of the World Wide Web. Your aim is to decide whether the English used in that domain is closer to British English or American English (or neither), and to find English terminology specific to your chosen domain. You should choose a national WWW domain; collect and cleanse part of a "data warehouse" or representative offline dataset; analyse the dataset to identify significant terms specific to and/or characteristic of this domain; and produce a report on your task including a description of your methods and results. This exercise will give you practical experience of applying the CRISP-DM methodology; your report should include a section for each of the CRISP-DM stages.

The term "domain" is ambiguous: a "domain" can mean a general subject area; but WWW URLs are generally divided into "domains" corresponding to administrative units, at top level mainly national domains, see http://en.wikipedia.org/wiki/List_of_Internet_TLDs For example, URLs ending **.uk** are websites in the United Kingdom, and English text on these websites represents British English; URLs ending **.us** are websites in the United States of America, and English text on these websites represents American English; URLs ending **.gm** are in the Gambia, and English text on these websites represents Gambian English, which may be similar to British or American English, and/or it may use some terms which differ from British and/or American English. For this assignment, you will use the URL sense of "domain", as this makes it easy to identify and collect web-pages from a given domain. You must choose a national "Top Level Domain", and claim this by emailing eric@comp.leeds.ac.uk . Each student must investigate a different domain, so that the class as a whole will cover a wide range.
Do not choose .uk or .us – these domains will be used in examples shown in lectures, and I will provide .uk and .usa datasets as **gold standards** against which you can compare and contrast your own domain data. Also, **do not choose .com, .edu, .org, .tv, .eu** or similar domains which are NOT national domains, as they are not restricted to one national variety of English. You should choose a domain where English is one of the native languages, or where English is widely learnt and used as a second or foreign language; this probably includes most countries of the world.

The CRISP-DM methodology is designed for data mining consultants providing a service for a client or customer. For this coursework exercise, you are the consultant. I represent the client: ICE, the International Corpus of English research consortium. For over 10 years, ICE has been trying to collect texts representing different national varieties of English, to use in comparisons with British and American English, and to identify specialised national terminology and other language features. So far they have only managed to collect text datasets or **corpuses** for a few national Englishes, because collecting texts by traditional means is slow and expensive. ICE is interested in trying to use the **web as corpus** approach instead of collecting traditional documents. Your consultancy report will show ICE whether and how the WWW can be used as a data-source to compare national varieties of English, and to identify domain-specific English terms.

You are to apply the CRISP-DM methodology to this knowledge management task:

- **Business Understanding** involves analysis of the aims of the "client", represented here by myself: the client wants to see how domain-specific WWW texts can be collected and used, to compare with British and American corpuses, and to identify terms specific to a national variety of English.
- **Data Understanding** includes collecting, describing and exploring the data for your domain. To collect the data, you can use a web-search engine such as Google or Yahoo, restricting the search to English-language pages; or existing web-as-corpus tools made available by Kilgarriff (SketchEngine), Pomikálek (WWWBootCat),

Roberts (Java JBootCat), Baroni (Perl BootCat), RDUES (WebCorp), Bernhard (Python WebCorpusBuilder), or Sharoff (Internet Corpus Query engine); or build your own software on top of Google/Yahoo API. You will need to register your own Google API key to use any systems which accees Google API, eg BootCat, WWW-BootCat etc. You should aim to collect a corpus of at least 200,000 words of English text representing your domain (and you should also record the source URLs)

- **Data Preparation** includes cleansing and "filtering" the websites, to extract the English text. Some web as corpus tools like Bootcat and SketchEngine include features to do this for you; for an additional challenge, you could code this yourself in Python or Java.

- **Modelling** is the Computational Modelling core of this exercise. A language model encapsulates characteristics of the language to be modelled. It involves analysis of your corpus to identify significant similarities and differences between your corpus and the .**uk** and .**us** gold standards. You need to gather evidence showing whether your corpus is "closer" to the .uk or the .usa sample; or maybe it is significantly different from both, if you find terms which are significantly common in your corpus, but not in British or American English as represented in the corpora provided. This can be done simply by examining and comparing frequent-word lists for the three corpora, and noting differences; or by using text analysis tools such as those provided with WebBootCat; or you may implement your own comparison tools.

- **Evaluation** includes assessment of the results with respect to business understanding: have you found evidence to reach a conclusion, and can you suggest explanations for the differences you have found?

- **Deployment** in this exercise involves writing a final report of the exercise. It includes reporting on all of the previous stages. You should also review the process you have gone through, and consider whether and how your approach could be improved

DELIVERABLES
Work through the CRISP-DM methodology, recording your work on each stage.

Write a short report entitled " **Which English dominates the World Wide Web in XX: British or American? (or neither?)**" (no longer than 6 pages; replace **XX** with your chosen domain). Your report format must conform to the NIPS2005 formatting instructions specified in http://www.comp.leeds.ac.uk/db32/documents/format.rtf - however note that your page limit is 6 pages, not 8 as required for NIPS2005. An example of a report formatted according to this template is my "CHEAT" approach to the MorphoChallenge2005 contest, see http://www.comp.leeds.ac.uk/db32/documents/cheat.doc The easiest way to ensure your report meets these formatting guidelines is to "cut out" the existing content from one of these two examples, and replace with your own content. In the report, include a subsection for each of the stages in the CRISP-DM guidelines, reporting how you approached each stage; sections need not all be the same length (eg the Business Understanding section should be concise), and you may decide to merge Data Understanding and Data Preparation as they are interlinked:

Business Understanding: what you think your client wants from this exercise

Data Understanding and Preparation: data sources and software tools you used, or developed yourself

Modelling: techniques/tools used for comparisons, and results, eg what types of correlations and/or domain-specific terms you found; **please include a summary of results which could be cut and pasted into a Research Report collating all results into an overview of the whole WWW, jointly authored by lecturer and students.**

Evaluation: whether the exercise was a success; and other knowledge you discovered about the dataset

Deployment: review of the exercise

Remember to attach the report to a standard Header sheet with your name etc; and submit via CSO coursework letter-box. **ALSO** submit an email (from your mailbox at comp.leeds.ac.uk or leeds.ac.uk, **not** an external account I cannot identify) to eric@comp.leeds.ac.uk with attachments: (i) your report.doc (MS-Word doc file), (ii) your corpus (linux text file), (iii) complete wordlist extracted from the corpus (linux text file), (iv) list of source URLs (linux text file, one URL per line). **NOTE**: only the Report should be in MS-Word format; all the other files must be plain text, viewable with linux text-editors eg vi or vim.

Marking scheme: Your report will be graded using the standard Informatics MSc Report Marking Form, see http://www.comp.leeds.ac.uk/db32/documents/report_mark_sheet.doc

This cw contributes 30% of your overall CMD grade.

Useful resources:

Atwell E. 2006. CMD Learning and Teaching Resources on Computational Modelling
http://www.comp.leeds.ac.uk/cmd/

Atwell E. 2006. DB32 Learning and Teaching Resources on Technologies for Knowledge
Management http://www.comp.leeds.ac.uk/db32/

Atwell E and Roberts A. 2006.Combinatory Hybrid Elementary Analysis of Text.
MorphoChallenge2005 Workshop proceedings, Venice.
http://www.comp.leeds.ac.uk/db32/documents/cheat.doc

Baroni M and Bernardini S. 2004. BootCaT: Bootstrapping corpora and terms from the web.
Proceedings of LREC 2004, Lisbon.
http://sslmit.unibo.it/~baroni/publications/lrec2004/bootcat_lrec_2004.pdf

Baroni M. 2005. BootCaT: Perl Utilities to Bootstrap Corpora and Terms from the Web:
http://sslmit.unibo.it/~baroni/bootcat.html

Baroni M et al. 2006. Wacky website for Web-as-Corpus kool ynitiative:
http://wacky.sslmit.unibo.it/

Bernhard D. 2006. Python Web Corpus Builder.
http://www.comp.leeds.ac.uk/db32/cw/web_corpus_builder/README

Bernhard D. 2005. Term Clouds demo website:
http://www-timc.imag.fr/Delphine.Bernhard/index.php?rub=3&subrub=termclouds&lang=en

Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C and Wirth R. 2000.
CRISP-DM step-by-step data mining guide http://www.crisp-dm.org/CRISPWP-0800.pdf

Hippocampus D. 2005. Formatting Instructions for NIPS*2005 (template for db32cw reports).
http://www.comp.leeds.ac.uk/db32/documents/format.rtf

Kilgarriff A. 2006. SketchEngine http://www.sketchengine.co.uk/

Nelson G. 2006. ICE: International Corpus of English website http://www.ucl.ac.uk/english-
usage/ice/

Pomikálek J. 2006. BootCaT WWW version: http://corpora.fi.muni.cz/bootcat/

RDUES. 2005. WebCorp website: advanced search:
http://www.webcorp.org.uk/wcadvanced.html
publications: http://www.webcorp.org.uk/publications.html

Roberts A. 2006 JBootCat http://www.andy-roberts.net/software/jbootcat/index.html   Also
registered with Google: http://code.google.com/p/jbootcat/

Sharoff S. 2006. A query to Internet corpora: website: http://corpus.leeds.ac.uk/internet.html

Appendix B: Computational Modelling second coursework assignment
(lecturer: Eric Atwell, School of Computing, University of Leeds)

Explaining your computational modelling to an interdisciplinary audience

Your previous CMD coursework assignment was an exercise in computational modelling of data from a specific World Wide Web national domain of your choice, to address the question: "Which English dominates the World Wide Web in your national domain: British or American?" I asked you to conform to the standard CRISP-DM structure and sections in your Report, as this standard is widely used and understood in computational data modelling and data mining research.

A challenge for Computing students and researchers working on interdisciplinary projects is how to communicate and report on their work in a way that is accessible to others with different backgrounds. For example, Cognitive Systems researchers may have to describe their work to psychologists, philosophers and linguists; Health Informatics researchers have to explain their computational modelling to health and medical practitioners; Bioinformatics researchers have to report on their implementations and results in a way which biologists will appreciate.

For your final CMD coursework, I want you to gain experience of writing for an interdisciplinary readership. Your challenge is to write a paper to submit to an interdisciplinary journal, reporting on your coursework2 findings for your chosen national domain, extending these to include comparative results from several other students for other countries in the region.   For example, if you studied English in the Sudan, you could report on English in Africa by comparing your results to those of students studying other African countries; or if you studied the English of the Cook Islands, your report could compare your results to findings for other Pacific Rim domains; or if you studied the English in Iranian websites, you could compare this to English used in other nearby countries such as Kuwait, India, Russia. You should aim to compare your results with several others; I don't want to set strict limits on how many other countries but I would expect to see between three and ten other countries compared.

Your first step is to choose a journal which might publish a paper including your domain, for example a journal dealing with language and cultural studies in a region, or variation in English, or applications of computational modelling in computing/ web technologies/ computing education etc, or even an Information Systems or Informatics journal which publishes papers on interdisciplinary applications; for example, visit the Library and browse (literally!) current journals to see what sort of topics are covered eg in the Languages section; or else search for online websites for relevant journals, eg:

English Today http://journals.cambridge.org/action/displayJournal?jid=ENG

European Journal of American Culture
http://www.intellectbooks.co.uk/journals.appx.php?issn=14660407

International Journal of Asian Studies
http://journals.cambridge.org/action/displayJournal?jid=ASI

International Journal of Middle East Studies
http://journals.cambridge.org/action/displayJournal?jid=MES

Journal of Arab & Muslim Media Research
http://www.intellectbooks.co.uk/journals.php?issn=17519411

Journal of Latin American Studies http://journals.cambridge.org/action/displayJournal?jid=LAS

Journal of Modern African Studies
http://journals.cambridge.org/action/displayJournal?jid=MOA

Journal of Southeast Asian Studies
http://journals.cambridge.org/action/displayJournal?jid=SEA

Nordic Journal of Linguistics http://journals.cambridge.org/action/displayJournal?jid=NJL

Studies in Second Language Acquisition
http://journals.cambridge.org/action/displayJournal?jid=SLA

Language Variation and Change http://journals.cambridge.org/action/displayJournal?jid=LVC

English Language Teaching: ELT Journal http://eltj.oxfordjournals.org/

International Journal of Francophone Studies
http://www.intellectbooks.co.uk/journals.appx.php?issn=13682679

International Journal of Lexicography http://ijl.oxfordjournals.org/

Iran Journal http://www.bips.ac.uk/journal.html

Journal of Applied Ontology http://www.applied-ontology.org/

Natural Language Engineering http://journals.cambridge.org/action/displayJournal?jid=NLE

The Computer Journal http://comjnl.oxfordjournals.org/

International Journal of Intercultural Information Management
http://www.inderscience.com/browse/index.php

International Journal of Knowledge Management Studies
http://www.inderscience.com/browse/index.php

International Journal of Internet Marketing and Advertising
http://www.inderscience.com/browse/index.php

International Journal of Foresight and Innovation Policy
http://www.inderscience.com/browse/index.php

I J of Web Based Communities http://www.inderscience.com/browse/index.php?journalID=50

I J of Web Engineering and Technology
https://www.inderscience.com/browse/index.php?journalID=48

I J of Innovation & Regional Development
http://www.inderscience.com/browse/index.php?journalCODE=ijird

I J of Business Intelligence and Data Mining http://www.inderscience.com/browse/index.php

I J of Information and Communication Technology
http://www.inderscience.com/browse/index.php

I J of Teaching and Case Studies
http://www.inderscience.com/browse/index.php?journalID=196

Having chosen an appropriate journal, you should study the guidance for authors, and revise and extend your previous coursework report to

(i)   suit the readership and aims of the journal – you should aim to explain your survey to academics who may be experts in the domain but not in computational modelling;

(ii)   (ii) extend your survey by comparing your results with other students' findings for neighbouring or related countries – aim for three to ten other findings;

(iii)   (iii) meet the format and length requirements of the journal – typically aim for about 10 pages, but if the journal specifies a lot more you should discuss this with me.

DELIVERABLES

You must hand in a hard copy of your paper, conforming to Journal format and length guidelines.
Remember to attach the report to a standard Header sheet with your name etc; and submit via CSO coursework letter-box. **ALSO** submit an email (from your mailbox at comp.leeds.ac.uk or leeds.ac.uk, **not** an external account I cannot identify) to eric@comp.leeds.ac.uk stating the NAME and URL of the journal you have chosen to submit to; and with an attachment: your report.doc (MS-Word doc file, unless the journal specifies another file format),

Marking scheme: Your report will be graded using the standard Informatics MSc Report Marking Form, see http://www.comp.leeds.ac.uk/cmd/documents/report_mark_sheet.doc

This cw contributes 50% of your overall CMD grade.