

University of Glasgow

Cultural Aspects to Google Searching

James Currall & Michael Moss

May 2008

Introduction

“Does Google provide the same answers to all users of a particular language (phonic community) irrespective of where they are in the world or are the results returned from a search highly localised?”

Our experiments so far suggest that neither one nor the other is universally true and at which end of the scale a particular search lies depends very much on what the search is about. This raises a further question as to:

“whether or not there are cultural biases in the way that Google indexes the web, decides on relevance and orders the results”

Jean-Noël Jeanneney in his well publicized book *Google and the Myth of Universal Knowledge*¹ condemned Google as a vehicle if not for American imperialism then at least for an Anglophone world that has important ramifications in Europe that no longer has a common *lingua Franca*. There can be little disagreement that Google, like any search engine, is not neutral, in just the same way, as it is now generally accepted, that neither a catalogue nor the contents of a library are neutral. The interesting question is as to whether Google, in selectively searching the global content of the Internet, is more or less culturally neutral than a physical archive, library or museum.

Whilst initially interested in this matter in relation to variations **within** different culturally diverse Anglophone territories, we see value in the exploration being carried into other phonic communities, as a first step to looking at the differences **between** phonic communities.

In order to explore these issues, we need to involve people from as many locations in different cultures around the world as possible and if those people also have an interest in cultural aspects of technology, then so much the better.

An Experiment for CATaC'08

The organisers of CaTAC'08 have generously given us the opportunity to involve you, the participants of CaTAC'08, in a information gathering exercise prior to the conference and then involve us all in a workshop to explore some of the dimensions of cultural specificity, using the information that we have jointly gathered. We are excited by the prospect of this type of interaction, are grateful for the opportunity and hope that you will take the opportunity to be involved.

The remaining sections provide further details of our plans and how you can take part. If we have left anything unclear or you wish further information, you can contact us at the address in the final section.

¹ Jean-Noël Jeanneney, *Google and the Myth of Universal Knowledge : a view from Europe*, translated by Teresa Lavender Fagan, The University of Chicago Press, Chicago and London, 2007. ISBN 0-226-39577-4, originally published as *Quand Google défie l'Europe: plaidoyer pour un sursaut*, Mille et une Nuits, Paris, 2005. ISBN 2-84205-912-3.

The searches

Our work so far, with a handful of colleagues around the world, has suggested that interesting results might be expected from the 14 topic areas that are listed below. Some of these will show divergence whilst others should show a remarkable degree of convergence. In our session we wish to engage with the conference participants on the implications of a range of findings within global and specific cultural contexts.

In order to capture information from a diverse range of locations and enable you all to engage with our workshop session we would like to invite you to conduct searches within these areas. Those of you from Anglophone territories (broadly described) should carry out searches from amongst those indicated under the topics below. For strict comparability, quotation marks and capitalisation should be as indicated (note that the last search has two quoted phrases, **in a single search**).

We would not presume to try to translate these terms into other languages for other phonic communities, not least because of the nuancing of language which is not obvious to non-native speakers. We would suggest that the phonic communities (Spanish, French, etc.) agree amongst themselves (perhaps facilitated through the CaTAC'08 forums) on appropriate terms within the topic areas, which in some cases might be the same concept and in others such as "literature/writers/poets" it might be appropriate to search for something altogether different.

1) art/artists

"impressionism"

2) literature/writers/poets

"Shakespeare"

"Robert Burns"

3) music/musicians/composers

"jazz"

4) religion/beliefs

"Christianity"

5) ethics/philosophy & science/technology

"genetically modified crops"

6) language

"minority language"

7) history/historical events/historical figures

"colonial history"

"Battle of the Somme"

8) race/ethnicity/racism

"Islamaphobia"

"racial minority"

9) conflicts

"terrorism"

"Afghanistan"

10) nationality/nation/nationalism

"Tibet"

11) tradition

"traditional dance"

12) current affairs

“credit crunch”

13) leisure pursuits

“birdwatching”

14) globalisation

“Johnnie Walker”

“Tony Blair”

“John Howard” “Prime Minister”

We will be supplementing this work with comparisons taken from searches of the catalogues of National Libraries in a number of Anglophone territories.

Timing

In order to minimise the variation resulting from timing, we suggest that the searches are carried out between **12pm and 1pm GMT on Monday 2nd June 2008** or within a few hours either way of that time. In selecting any time, we realise that we are going to choose a time that is impractical for some and for that we apologise, but no time is universally convenient and that time covers a wide region.

Important General Points

The points below are intended to reduce misunderstandings that might arise from different interpretations of the outline above.

Getting Google to Search the Correct Corpus

In most territories, Google allows you to search either ‘**the web**’ or ‘**pages from ...**’ where ‘...’ is your region. It is important that everyone does their search on ‘**the web**’. Such a search would seem to imply that we are all searching the same corpus, however we can already demonstrate that this is not the case. It is the differences that exist between searches that on the face of it one might expect to be similar that are the focus of our attention. We would naturally expect searches on pages located in different territories to be substantially different and are not therefore interested in ‘**pages from ...**’ searches.



The screenshot shows the Google search interface. At the top, there are navigation links: Web, Images, Maps, News, Shopping, Mail, and more. A 'Sign in' link is on the right. The Google logo is on the left, and the search input field contains the text "bird watching". To the right of the input field is a 'Search' button and links for 'Advanced Search' and 'Preferences'. Below the input field, there are radio buttons for 'the web' (selected) and 'pages from the UK'. At the bottom of the search results area, it says 'Web Results 1 - 10 of about 4,640,000 for "bird watching" [definition]. (0.24 seconds)'.

Country-Specific Google URLs

Google has search pages in different Internet Domains, e.g. ‘www.google.co.uk’, ‘www.google.fr’, etc.

 <http://www.google.co.uk/search?l>

It is important to note that in some territories, when you type a Google address into your browser you will be redirected in one of three ways:-

- From 'www.google.com' to your local search page (e.g. for us 'www.google.co.uk')
- To 'www.google.com' no matter what you type.
- To whatever you type.

When conducting searches it would be helpful if you could note what happens in this respect. We would prefer the search page of your region (in our case 'www.google.co.uk'). We have also seen some browser specificity in behaviour and so it would also be helpful if you also note which browser software you are using.

Capturing the results

It would be most useful if the result of each of your searches could be captured to a file (preferably a PDF). This means that you could send us your PDF files ahead of CaTAC'08.

For those of you that do not have Adobe Acrobat PDF creation software or a browser that offers to save or 'print' PDFs, there are other packages that can do this. For the Windows environment there is PDFCreator which is available free from the following URL:-

- <http://sourceforge.net/projects/pdfcreator/>

This software installs as just another printer and so you print to 'PDFCreator' rather than to a physical printer.

It would also be helpful if you could name your files with the following convention (it will help us manage them on our machines:-):

- `[territorial suffix] [search terms].pdf`

For example we might save them thus:-

 UK Admiral Lord Nelson.pdf	29 KB	Adobe Acrobat 7...	23/03/2008 20:58
 UK bird watching.pdf	45 KB	Adobe Acrobat 7...	23/03/2008 20:58
 UK colonial history.pdf	32 KB	Adobe Acrobat 7...	23/03/2008 20:54
 UK credit crunch.pdf	34 KB	Adobe Acrobat 7...	23/03/2008 20:55
 UK Iraq.pdf	33 KB	Adobe Acrobat 7...	23/03/2008 20:59
 UK John Howard Prime M...	44 KB	Adobe Acrobat 7...	23/03/2008 21:00
 UK search engine.pdf	30 KB	Adobe Acrobat 7...	23/03/2008 21:01
 UK terrorism.pdf	31 KB	Adobe Acrobat 7...	23/03/2008 20:57
 UK Tibet.pdf	33 KB	Adobe Acrobat 7...	23/03/2008 20:56
 UK Tony Blair.pdf	41 KB	Adobe Acrobat 7...	23/03/2008 20:59

Finally

We greatly appreciate your involvement, should you wish to participate in this, whether or not you wish to send us results of your searches. We are looking forward to a very lively session which will allow us to explore the privileging of information returned from search engines and discuss together the cultural implications of what we observe and ramifications for information services.

Our email address for any searches that you wish to send us is: j.currall@compserv.gla.ac.uk