The nationality of issues
Repurposing Google for Internet Research

DMI Summer School

Session 3

Facilitated by Esther Weltevrede

1. The nationality of issues
2. Google and location
3. DMI Google tools
4. Projects
5. The wiki

Studying the attention of national issues

- Teun van Dijk (1988)

Method for analyzing the international coverage of national news in the press

- Noortje Marres & Richard Rogers (2008)

Method for analyzing where an issue is happening and where it is discussed on the web

Question of the week: How can Google be repurposed for researching the nationality of issues on the web?

Structures of international news

News: discourse analysis

Case study: How the assassination of president-elect Bechir Gemayel of Lebanon (1982) is covered in the international press

- Coverage, frequency, relevance
- Headlines, type of article, section
- Thematic structure (actors, event, context, history, cause/effect)
- Topic clusters

Findings:

First world press more background topics, third world more about (international) reactions

Western press privileges western-related topics and actors

Subsuming the ground

Web: hyperlink analysis

Case study: How local realities of the Ferghana Valley, the Narmada Dams, and the BTC pipeline are put to use on the Web

Issue Crawler counts and weights hyperlinks

Actors and the articulation of issues derived from sites

Findings:

Issues are not articulated in more detail locally: “it cannot be maintained unconditionally that the closer one gets to the ground, the more one knows about what is going on there” (2008: 7)

Local realities tend to be presented as part of global controversies: “Local issues are mobilized, on the Web, to make global points” (2008:21)

Google and Internet research

Google repurposes news and hyperlinks in their search algorithms. Google creates ‘spheres’

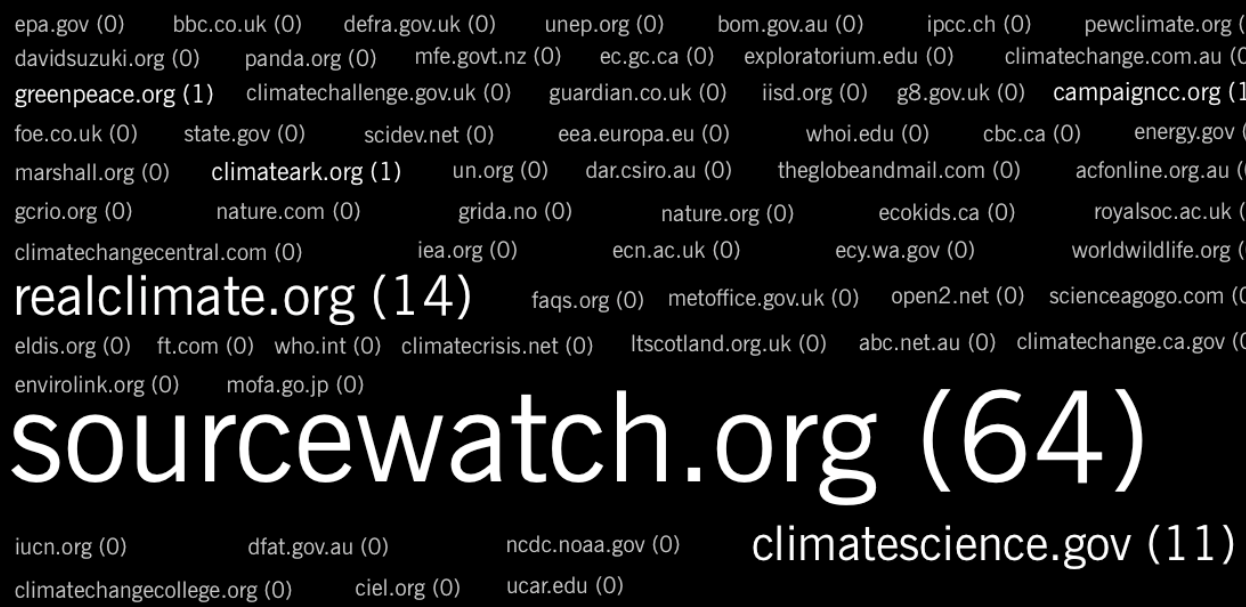
- News sphere (news.google.com)
- web sphere (google.com)
- blogosphere (blogsearch.google.com)

Not the habermasian ‘public sphere,’ but an engine demarcated source set, presented as ranked result pages

Previous DMI research: ‘source distance’ and ‘cross-spherical analysis’

Climate Change Sceptics on the Web (S. Fred Singer)

Research Question_To what extent are climate change 'skeptics' present in the climate change spaces on the Web?
Findings_There is distance between the skeptics and the top of the search engine returns.



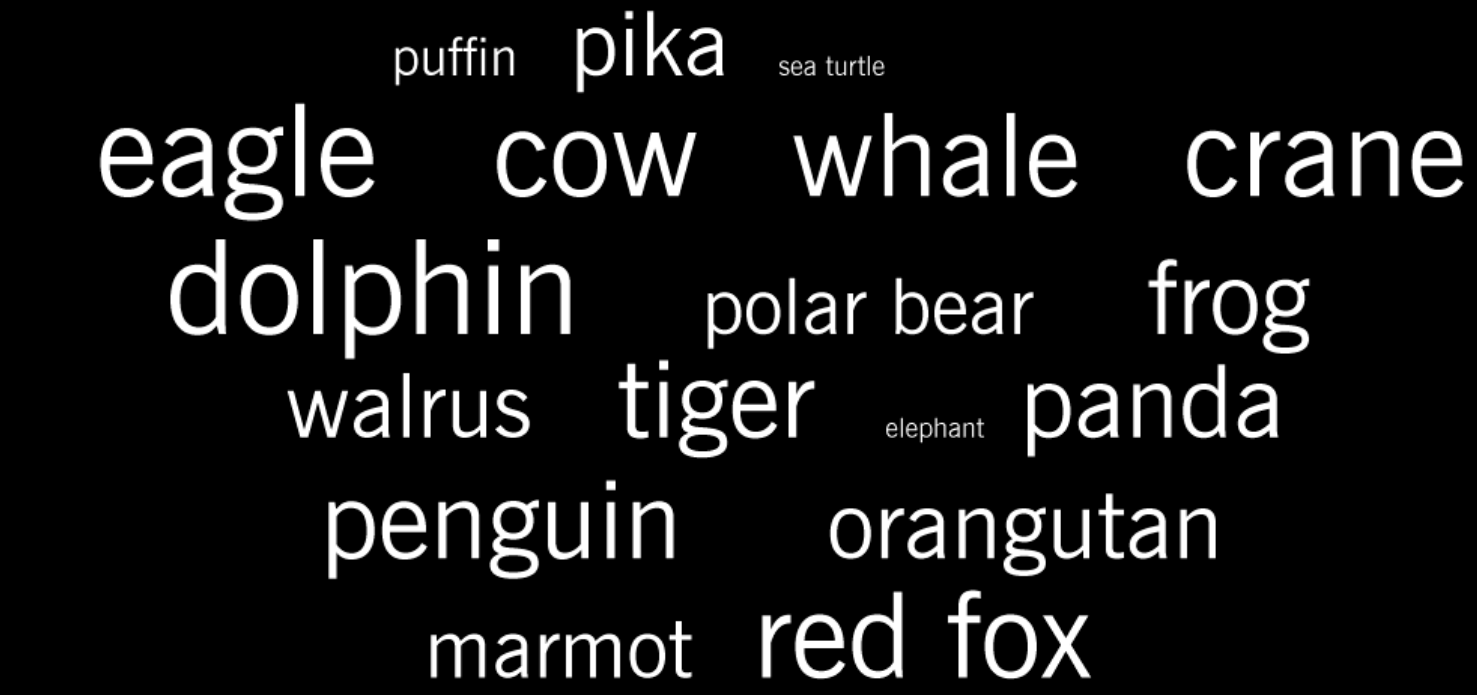
Source_google.com
Query_"Fred Singer"
Method_Search for query "Fred Singer" in top 100. Organized in order.
Tools_Google Scraper and Tag Cloud Generator
Date_30 July 2007

Product_of the Digital Methods Initiative, dmi.mediastudies.nl. **Analysis**_by Bram Nijhof, Richard Rogers and Laura van der Vlies. **Design**_Anne Helmond.

CLIMATE CHANGE SCEPTICS

Issue Animals Hierarchy on the Web (Google)

Research Question_For the issue of climate change, how prominent is each animal (in text and image)? Are there significant differences per 'sphere' (web, news, blogosphere) in the frequency with which each animal is referenced?
Findings_On the web, for a text query, results are distributed across all the animals not particularly favoring one issue animal.



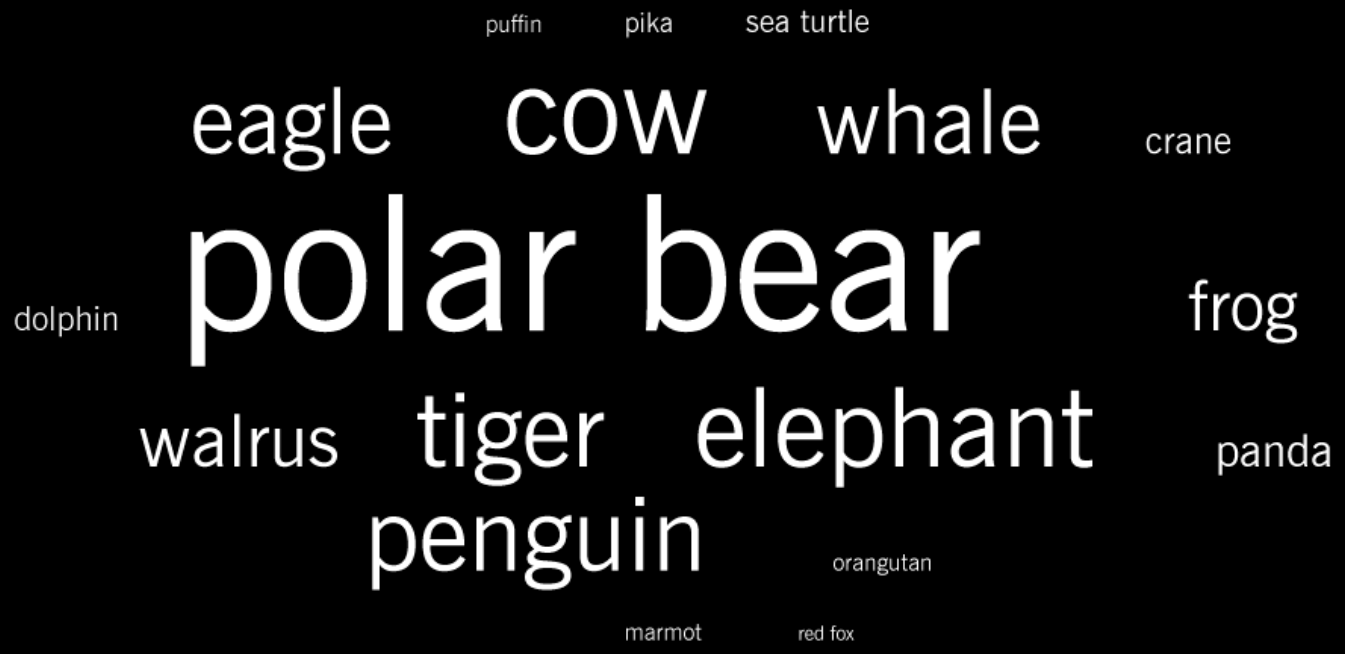
Source_www.google.com
Query_"climate change" + scrape top 100 results for "*animal x*"
Tools_Google Scraper, Compare Lists and Tag Cloud Generator
Date_15 July 2007

Product_of the Digital Methods Initiative, dmi.mediastudies.nl. **Analysis**_by Esther Weltevrede and Sabine Niederer. **Design**_by Esther Weltevrede, Sabine Niederer and Anne Helmond.

ISSUE ANIMALS

Issue Animals Hierarchy in the News (Google News)

Research Question_For the issue of climate change, how prominent is each animal (in text and image)? Are there significant differences per 'sphere' (web, news, blogosphere) in the frequency with which each animal is referenced?
Findings_In the news, for a text query, the polar bear is the animal most associated with climate change, followed by the cow.



Source_news.google.com
Query_"climate change" + *"animal x"*
Tools_Google News Scraper and Tag Cloud Generator
Date_15 July 2007

Product_of the Digital Methods Initiative, dmi.mediastudies.nl. **Analysis**_by Esther Weltevrede and Sabine Niederer. **Design**_by Esther Weltevrede, Sabine Niederer and Anne Helmond.

ISSUE ANIMALS

Issue Animals Hierarchy in the Blogosphere

Research Question_For the issue of climate change, how prominent is each animal (in text and image)? Are there significant differences per 'sphere' (web, news, blogosphere) in the frequency with which each animal is referenced?
Findings_In the blogosphere, for a text query, the polar bear is the animal most associated with climate change, followed by the cow.

Source_search.technorati.com
Query_"climate change" + "*animal x*"
Authority_a little
Tools_Technorati Scraper and Tag Cloud Generator
Date_17 July 2007

Product_of the Digital Methods Initiative, dmi.mediastudies.nl. **Analysis**_by Esther Weltevrede and Sabine Niederer. **Design**_by Esther Weltevrede, Sabine Niederer and Anne Helmond.

ISSUE ANIMALS

Google for national web studies

Most Google engines have a national version

http://www.google.com/
http://news.google.com/
http://blogsearch.google.com/ (the exception)

Visit Google's Site in Your Local Domain

www.google.ad Andorra
www.google.ae الامارات العربية المتحدة
www.google.com.af افغانستان
www.google.com.ag Antigua and Barbuda
www.google.com.ai Anguilla
www.google.am Հայաստան
www.google.it.ao Angola
www.google.com.ar Argentina
www.google.as American Samoa
www.google.at Österreich
www.google.com.au Australia
www.google.az Azərbaycan
www.google.ba Bosna i Hercegovina
www.google.com.bd বাংলাদেশ
www.google.be België
www.google.bg България
www.google.com.bh البحرين
www.google.bi Burundi
www.google.bj Bénin
www.google.com.bn Brunei
www.google.com.bo Bolivia
www.google.com.br Brasil
www.google.bs The Bahamas
www.google.co.bw Botswana
www.google.com.by Беларусь
www.google.com.bz Belize
www.google.ca Canada
www.google.cd Rep. Dem. du Congo
www.google.cf Centrafrique
www.google.cg Rep. du Congo
www.google.ch Schweiz
www.google.ci Cote D'Ivoire
www.google.co.ck Cook Islands
www.google.cl Chile
www.google.cn 中国
www.google.com.co Colombia
www.google.co.cr Costa Rica
www.google.com.cu Cuba
www.google.cz Česká republika
www.google.de Deutschland
www.google.dj Djibouti
www.google.dk Danmark
www.google.dm Dominica
www.google.com.do Rep. Dominicana
www.google.dz Algérie
www.google.com.ec Ecuador
www.google.ee Eesti
www.google.com.eg مصر
www.google.es España
www.google.com.et ኢትዮጵያ
www.google.fi Suomi
www.google.com.fj Fiji
www.google.fm Micronesia
www.google.fr France
www.google.ge საქართველო
www.google.gg Guernsey
www.google.com.gh Ghana
www.google.com.gi Gibraltar
www.google.gl Grønland
www.google.gm The Gambia
www.google.gp Guadeloupe
www.google.gr Ελλάδα
www.google.com.gt Guatemala
www.google.gy Guyana
www.google.com.hk 香港
www.google.hn Honduras
www.google.hr Hrvatska
www.google.ht Haïti
www.google.hu Magyarország
www.google.co.id Indonesia
www.google.ie Ireland
www.google.co.il ישראל
www.google.im Isle of Man
www.google.co.in India
www.google.is Ísland
www.google.it Italia
www.google.je Jersey
www.google.com.jm Jamaica
www.google.jo الأردن
www.google.co.jp 日本
www.google.co.ke Kenya
www.google.com.kh កម្ពុជា
www.google.ki Kiribati
www.google.kg Кыргызстан
www.google.co.kr 한국
www.google.com.kw الكويت
www.google.kz Қазақстан
www.google.la ລາວ
www.google.com.lb لبنان
www.google.li Liechtenstein
www.google.lk Sri Lanka
www.google.co.ls Lesotho
www.google.lt Lietuvos
www.google.lu Luxemburg
www.google.lv Latvija
www.google.com.ly ليبيا
www.google.co.ma Maroc
www.google.md Moldova
www.google.me Crna Gora
www.google.mg Madagasikara
www.google.mn Монгол улс
www.google.ms Montserrat
www.google.com.mt Malta
www.google.mu Mauritius
www.google.mv Maldives
www.google.mw Malawi
www.google.com.mx México
www.google.com.my Malaysia
www.google.co.mz Moçambique
www.google.com.na Namibia
www.google.com.nf Norfolk Island
www.google.com.ng Nigeria
www.google.com.ni Nicaragua
www.google.nl Nederland
www.google.no Norge
www.google.com.np नेपाल
www.google.nr Nauru
www.google.nu Niue
www.google.co.nz New Zealand
www.google.com.om عمان
www.google.com.pa Panamá
www.google.com.pe Perú
www.google.com.ph Pilipinas
www.google.com.pk Pakistan
www.google.pl Polska
www.google.pn Pitcairn Islands
www.google.com.pr Puerto Rico
www.google.pt Portugal
www.google.com.py Paraguay
www.google.com.qa قطر
www.google.ro România
www.google.ru Россия
www.google.rw Rwanda
www.google.com.sa السعودية
www.google.com.sb Solomon Islands
www.google.sc Seychelles
www.google.se Sverige
www.google.com.sg Singapore
www.google.sh Saint Helena
www.google.si Slovenija
www.google.sk Slovensko
www.google.com.sl Sierra Leone
www.google.sn Sénégal
www.google.sm San Marino
www.google.st São Tomé e Príncipe
www.google.com.sv El Salvador
www.google.co.th ประเทศไทย
www.google.com.tj Tajikistan
www.google.tk Tokelau
www.google.tl Timor-Leste
www.google.tm Türkmenistan
www.google.to Tonga
www.google.com.tr Türkiye
www.google.tt Trinidad and Tobago
www.google.com.tw 台灣
www.google.co.tz Tanzania
www.google.com.ua Україна
www.google.co.ug Uganda
www.google.co.uk UK
www.google.com.uy Uruguay
www.google.co.uz O'zbekiston
www.google.com.vc Saint Vincent and the Grenadines
www.google.co.ve Venezuela
www.google.vg British Virgin Islands
www.google.co.vi Virgin Islands
www.google.com.vn Việt Nam
www.google.vu Vanuatu
www.google.ws Samoa
www.google.rs Србија
www.google.co.za South Africa
www.google.co.zm Zambia
www.google.co.zw Zimbabwe

Things to take into account using Google for national web studies

- Digital divide
- Market share
- Language
- Availability in Google

Digital Divide Cartogram

Country size is proportional to level of **Internet usage**.

Data source: internetworldstats.com

Product of the Information Technology and International Cooperation Program.

Amsterdam, The Netherlands. April **05**

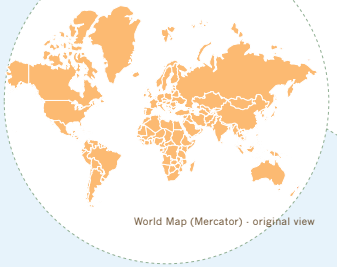
Analysis_Koen Martens, sonologic.nl.
Design_Marieke van Dijk, anderemedia.nl.

SSRC

Digital Divide Cartogram

Country size is inversely proportional to level of **Internet usage.**

Data source: internetworldstats.com



Product of the Information Technology and International Cooperation Program.

Amsterdam, The Netherlands. April **05**

Analysis_Koen Martens, sonologic.nl.
Design_Marieke van Dijk, anderemedia.nl.

SSRC

National market share

"Google matters most in Latvia, Lithuania, Hungary, Poland, Romania, Belgium, and the Netherlands. In each of these nation states Google controls more than 95 percent of the Web search market."

"Closely behind these leaders are Venezuela, Switzerland, Spain, Portugal, Italy, Germany, France, Finland, Denmark, Columbia, Chile, Brazil, Argentina, and the United Kingdom, all of which give Google between 90 and 95 percent of their Web search traffic.",

"Google is far behind the local competition in China, Hong Kong, Japan, Taiwan, South Korea, and Russia. Each of these nation-states grants Google less than 40 percent of the search market."

- Siva Vaidhyanathan, June 30, 2009

Also check: http://alexa.com/topsites/countries

Language

http://www.google.com/language_tools

Some other interesting Google tools

http://www.google.com/trends

http://www.google.com/insights/search/#

https://adwords.google.com/select/KeywordToolExternal

Availability in Google

http://www.google.com/support/insights//bin/answer.py?answer=99452

Project proposals

1. Cross-national issue analysis (compare source-distance, cross-spherical analysis)

– how local or global is an issue?

2. Search engine monitoring (compare Personalized search)

– How far along is customization on location? What is the difference in results between Google's different national webs (e.g. advanced region vs google.nl)

3. Customization on location versus Internet censorship

– How to turn Google into a tool for Internet censorship research?

4. Topology of issues per national web

– how can Google be used to tell which issues are prominent in a national web space?

After coffee

- Generatenational.net
- Project groups
- The Wiki

The national web according to Google

Generatenational.net

1. Cross-national issue analysis

– how local or global is an issue?

Choice of issue:

- national issue across national webs: an exercise in studying where is an issue happening and where is it discussed (e.g. iran election)
- international issue in national web spaces: an exercise in studying the attention an issue receives nationally, and how (e.g. climate change/global warming/global cooling)
- studying a controversy across national webs?

2. Search engine monitoring

- how far along is customization? how much influence does customization have on results?

- Compare to language results or personalization? Use proxies?

Background reading: http://wiki.digitalmethods.net/Dmi/PersonalizedSearch

Customization on location versus Internet censorship

- how does customization on location relate to censorship?
- how do top results for a query compare to what people actually search for in a country?

Trends data is based on 'successful queries,' which could be used as indicator for censorship

http://trends.google.com/websites?q=http%3A%2F%2Fwww.humum.net%2F&geo=all&date=all

Background reading:

http://wiki.digitalmethods.net/Dmi/DmiSummer09#Recommended_Readings

Topology of national issues per national web (over time?)

- Comparing the prominence and relevance of issues across national webs
- Does the local web privilege international or national issues?

User studies: - what do people search for in a national web?

Generatenational: - what issues are prominent in the top results of a national web?

The wiki

http://wiki.digitalmethods.net/Dmi/ThingsInternetResearchersShouldKnowAboutGoogle