

Issue Crawler Web Network Mapping Software and Allied Tools

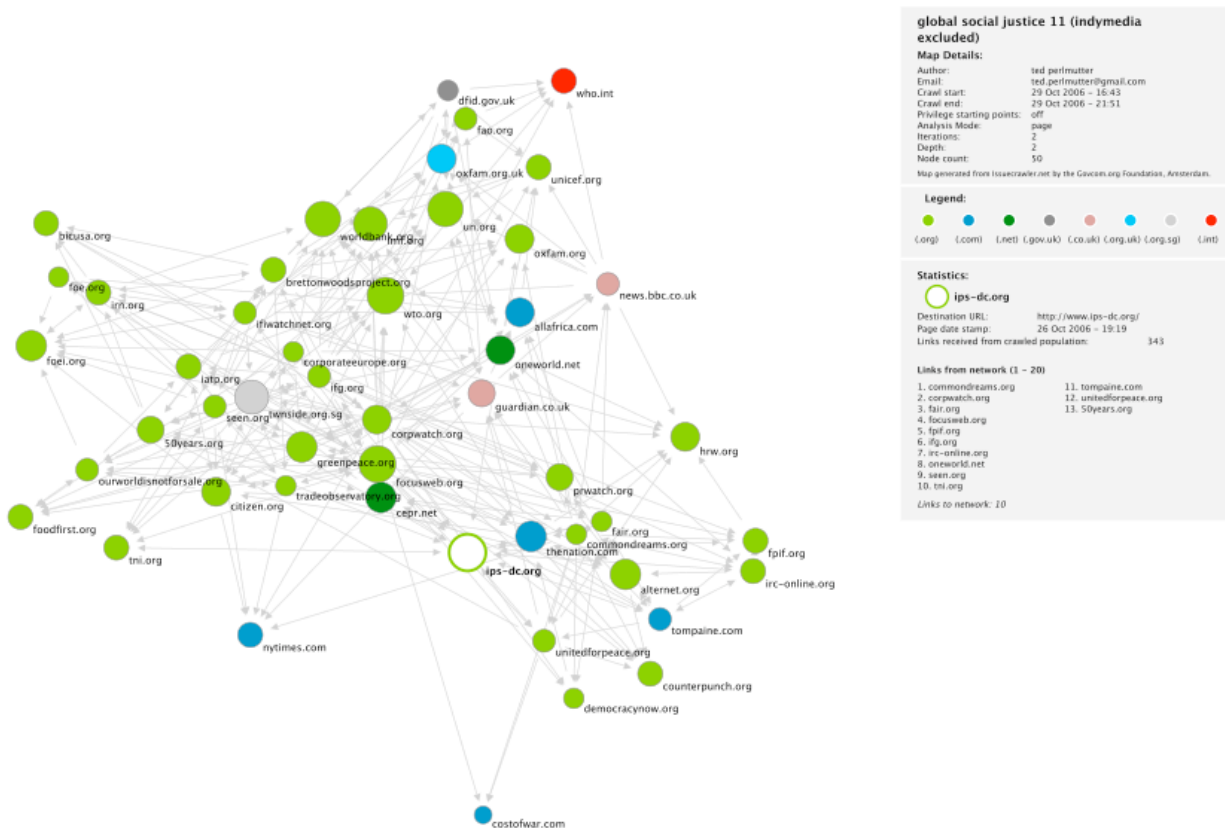
<http://www.issuecrawler.net>, <http://tools.issuecrawler.net>, <http://movies.issuecrawler.net> and <http://www.govcom.org>.

Issue Mapping Contextual Essay

Prepared by Richard Rogers, Director, Govcom.org Foundation, Amsterdam and Head of New Media, University of Amsterdam

Short Description of the Issue Crawler and Allied Tools

The Issue Crawler is Web network and visualization software that works in a browser. It consists of crawlers, databases, analysis engines, and visualization modules. The purpose of the software is to locate and visualize issue networks, clusters of organizations engaged in the same issue area. It relies on co-link analysis, a scientometric (or webmetric) technique based on citation analysis. The Issue Crawler also has allied tools, one of which, scrapeGoogle, allows one to query the issue network actors for substance, e.g., which organizations in the network make use of particular issue language.



Global Social Justice Network, Top 50 actors, Issuecrawler.net cluster 'map' output, October 2006.

For example, in the large 'social justice' network, which organizations refer to 'media justice' and how often? Thus, apart from 'mapping the network,' the software also provides indicators of the networked organizations' issue commitments. One also may monitor the growth or decline of the issue commitments over time, for the network as a whole, or for separate organizations.

Overall Objective of the Issue Crawler and Allied Tools and Their Theoretical Underpinnings

The objective of the larger project is to provide users with the capacity to map and analyze networks on the Web. The Issue Crawler software has been designed (with its default settings) to map issue networks, though users, through changing the settings, also may map social networks as well as establishment

networks. The notion of an issue network was developed in the mid-1970s by a writer from the right-of-center think tank, the American Enterprise Institute (AEI), as a means of describing and also warning against the rise in influence of clutches of NGOs, funders, think tanks and academics, developing powerful streams of thought as well as policy prior to the proper legislative procedure.¹ To the AEI writer, Hugh Hecla, issue networks were considered a threat to democracy. Govcom.org, through its software as well as its writings (including academic works), revived the concept of the issue network to point to its continued significance as a notion, without necessarily drawing from AEI's conservative school of thought. With the issue network, as well as other notions such as 'issue-fiction,' 'issue drift' and the life of an issue, Govcom.org contributes to the development of new political and social theory suitable for the contemporary period, with its emphasis on the importance of non-state actors (issue professionals), using new media. The Issue Crawler may be used as an instrument of empirical analysis as well as critique. It also may be used for event mapping: who's here, and who should be here?

Previously the U.S. notion of the 'issue network' was mentioned, but the cultural and geographical root of the project lies more in the Dutch tradition of debate mapping (*de sociale kaart* and *de brede maatschappelijke discussie*). In the early period of our work, we strove to map debates on the Web,² but as Noortje Marres opens her PhD dissertation, "we found issue networks instead."³



Narmada Dams Network rendered as Geographical Map, Issuegeographer output, redesigned, March 2005.

¹ Hugh Hecla, "Issue Networks and the Executive Establishment," in *The New American Political System*, American Enterprise Institute, Washington D.C., 1978, 87-124.

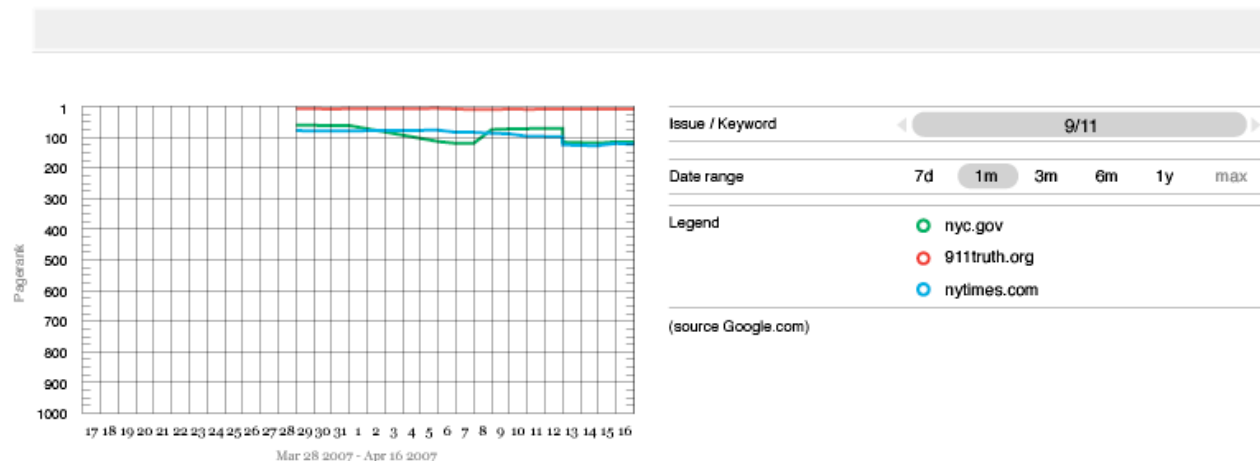
² See Richard Rogers, "Mapping Web Space with the Issue Crawler," unpublished ms., 2006, http://www.govcom.org/publications/full_list/issuecrawler_1oct06_final.pdf.

³ Noortje Marres, *No Issue, No Public: Democratic Deficits after the Displacement of Politics*, PhD Dissertation, University of Amsterdam, 2006.

The Issue Crawler also has a geographical visualization module, the Issue Geographer, which plots the actors' registered (whois) location to a geographical map, where the critical question concerns the distance between the actors and the issues on the ground. For example, in a study, rendered in a video, the question concerned the distance between the most recognized actors in the Narmada Dams issue network and the Narmada Valley itself.⁴

Recently Govcom.org developed the IssueDramaturg, which captures the prodigious climbs and precipitous drops in an actor's Google PageRank per issue query. The IssueDramaturg completes a project, conceived at the beginning of Govcom.org – the Preferred Placement period, where the drama surrounding top ten search engine placement was first sketched.⁵ In this novel technique it is determined whether an organization is to be found in Google returns for a particular issue query. For example, when searching for “radio consolidation,” where does futureofmusic.org appear in Google results? With the scheduler, one may watch rankings per issue over time. With site comparison, one may compare two or more sites per issue, also over time. For example, for a 9/11 query, the 9/11 Truth Movement remains stable at the top of Google returns, whilst New York City's official governmental Website as well as the *New York Times* decline below a PageRank of 100, a user's maximum default setting for returns per page.

Issue Dramaturg



IssueDramaturg by Govcom.org, April 2007, <http://issuedramaturg.issuecrawler.net>

To date Govcom.org's analyses has relied on text (as well as other media) published on the Internet, and its 'hypertextual apparatus,' as opposed to read text, and its cognitive or hermeneutic interpretations. In other words much of the work has had to do not with audience or reception, but with impact and resonance. The theoretical roots of the work lie in actor-network theory (ANT), and especially ANT's critique of 'diffusionist' models (sources and recipients). In circulation theory any actor or entity's capacities and agency are derived by its enrolment in a network. A practical example of the difference between diffusionist and circulationist theoretical commitments concerns how a researcher would approach Web presence. Web presence could be thought of in terms of a site's content and design – what is on it, how it is laid out, which features are incorporated. One 'manages' presence by updating and redesigning, having learned from the 'audience' of surfers, e.g., their paths taken through sites, their hit and visit counts, their downloads. Especially since the development of Google's PageRank and, later, Web 2.0's ratings culture, presence may be viewed as less controllable or manageable by the site owner. With Google a site appears in its returns by virtue of other sites' linking and pointing to it. With ratings culture, a site appears in

⁴ Govcom.org, *The Places of Issues: IssueCrawler Back-end Movie*, shown at "Making Things Public: Atmospheres of Democracy," curated by Peter Weibel and Bruno Latour, ZKM, Karlsruhe, 2005, <http://movies.issuecrawler.net>.

⁵ See Richard Rogers (ed.), *Preferred Placement: Knowledge Politics on the Web*, Jan van Eyck Editions, Maastricht, 2000.

aggregation and recommendation devices (e.g., del.icio.us) by virtue of other users' annotations and tags. Thus, the circulationist, when considering presence, would define a site and its contents through how it is indexed, linked, referenced, syndicated and tagged, and not by its content, freshness, design and features. The extent to which the site is referenced – or enrolled – by networks defines its presence. Finding the site, which is one means to think of a site's capacities and agency, is determined by the network.

More specifically, a site's presence steadies and climbs through references from others, especially from those sites that are themselves 'highly authoritative.' Is the site well-linked to, well-blogged, well-tagged, and by whom? 'Authority' online, as measured by everyday Web devices like Google, Technorati, Digg and others, begins with those simple starting points.

Keeping the basic insight of network-authored presence as well as authority in mind, Govcom.org has developed a project called 'advanced Web metrics.' It moves beyond the Websphere-only networks to other spaces, e.g., the news space, the blogosphere and the folksonomic.

New Project: Advanced Web Metrics

The advanced Web metrics project builds upon existing indexing and recommendation tools to provide indications of actor impact and resonance. Should an organization blog about issues, one may measure the extent to which the organization's particular postings resonate in the spaces in the blogosphere where the issue is discussed. To begin, one may measure the resonance of an actor per issue in the blogosphere. Building upon Technorati as well as Govcom.org's issuecrawler.net scrapers, this tool first ascertains an organization's issue blogging commitments by scraping all blog posts and tags, and subsequently finds whether those organization's issue blog postings are resonating in the blogosphere. For example, Public Knowledge blogs mostly about net neutrality and copyright. Does the blogosphere associate Public Knowledge with those issues or others? Subsequently one may determine the organization's relative resonance per issue in the blogosphere. Building upon Technorati, the tool shows a ratio of all issue postings to an organization's association with the issue postings. For example, of the 102 blog postings on "spectrum reform," how many mention Public Knowledge?

Net Neutrality

(186) Copyright (124) Fair

Use (77) DRM (63) Intellectual Property

(59) Broadcast Flag (58) FCC (37) Information Policy

(31) Broadband (26) WIPO Broadcasters Treaty (25) Piracy (22) Spectrum Reform (20) Municipal Wi-Fi (18) Orphan Works (17) Analog Hole

(15) Government Mandates (14) Copyright Modernization Act of 2006

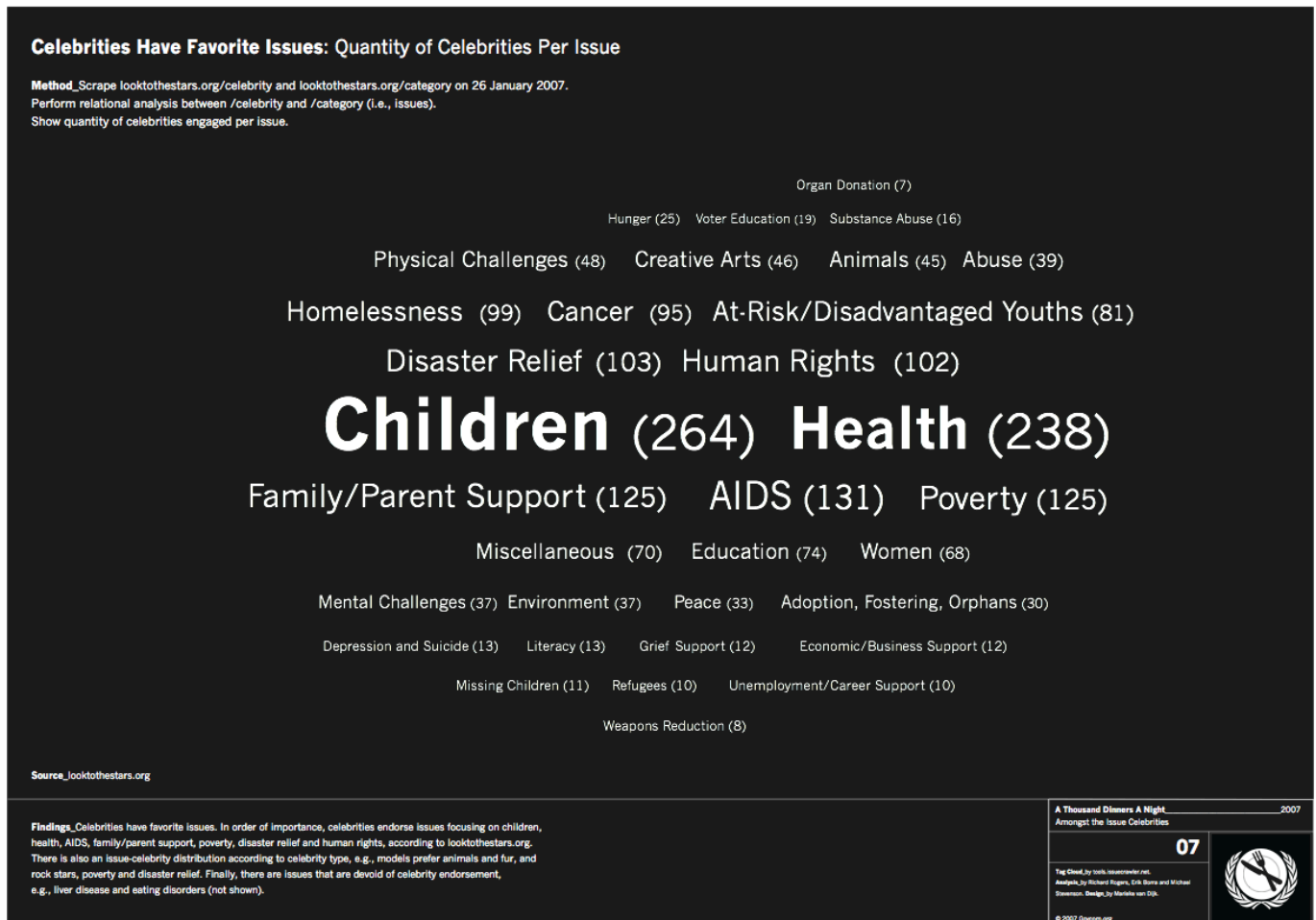
(13) Public Knowledge (12) Public Domain (10) Trademark (9) Video

Franchise (8) Patent (8) Open Access (7) Open Standards (6) P2P

(6) Internet Protocol (5) DTV (4) HD Radio (3) IP3 (2) HDTV (2)

Public Knowledge's Issue Resonance in the Blogosphere (Tag Cloud)

Govcom.org recently embarked on a ‘tagology’ project as a part of the advanced Web metrics undertaking, with its Del.icio.us organization tag cloud generator per issue. Building upon Del.icio.us, this tool shows, in a tag cloud, which URLs (organization domains like epic.org) are referred to per issue area (e.g., privacy).



Quantity of Celebrities per Issue, Govcom.org, February 2007.

Appendix: Project History

The project’s origins lie in an observation I made in 1996 when I was asked to write a newspaper article on climate change in the run-up to Kyoto. (The story is told in the foreword to *Information Politics on the Web*, MIT Press, 2004.) I typed climate change into AltaVista, and browsed through the returns. Looking to understand what at the time was called the ‘value of information’ or a sense of the reputation of sites, I noticed that Websites link selectively, as opposed to capriciously. In 1998 together with students in the Department of Science Dynamics at the University of Amsterdam and in Computer-Related Design at the Royal College of Art (where Josh On also studied at the time), we drew nodes and links on a chalkboard with colored chalk. The project began in earnest when I was named Design and Media Research Fellow at the Jan van Eyck Academy, Maastricht, and engaged the services of former students and colleagues from Amsterdam and London: Stephanie Hankey, Noortje Marres, Ian Morris and Alex Wilkie. Nick Durrant and Andres Zelman also participated, as did Peter Bilak and Anja Lutz. We continued to map links, also developing the NetLocator (aka the Depluralsing Engine), the forerunner to the Issue Crawler. It was dubbed the depluralsing engine because hyperlinking patterns showed not the hopeful neo-pluralistic potential of the Web, but rather depluralsing tendencies. Certain sites gain in stature (and in placement in engine returns) by virtue of the quantity and type of links they receive. Others become buried.

In the Preferred Placement period at the Jan van Eyck we sought to describe the delicate sociality of making or not making a link as hyperlink diplomacy. We also argued that the links show a rather normal 'politics of association'. The visual language for linking styles (totem, extracurricular, 'fully transdiscursive', etc) as well as linking types (cordial, critical and aspirational) were also developed at this time. Govcom.org, the name, was coined, when we noticed that the main actors involved in issues (when performing hyperlink analysis) were '.gov's', '.com's' and '.org's'.

The hand-made map we inserted in the Preferred Placement was shown to Jonathan Peizer of the Soros Internet Program, whereupon he asked whether we could build a machine to generate such a map. The first version of the Issue Crawler project (2001) was born of a collaboration between Govcom.org and OneWorld.net, where we worked intensively with David Heath and Suzi Wells. We also held workshops in Budapest, at C3, the first four in "The Life of Issues" series. During this time Govcom.org also developed, with Marieke van Dijk and Auke Touwslager as well as project collaborators from Cluj-Napoca and London, the Viagratool, the Issue Ticker (infoid.org), the Issue Barometer as well as the Election Issue Tracker (during the Pim Fortuyn period), with the support of Infodrome, the Dutch governmental information society initiative. The Issue Ticker has been exhibited at ZKM, Karlsruhe (together with Issue Crawler maps and the Issue Crawler Back-end Movie in "Making Things Public: Atmospheres of Democracy") as well as La Casa Encendida, Madrid, in "Derivatives: New Finance Art Visions." Issue Crawler and other Govcom.org maps have been exhibited at such events as the launch of Psiphon (Toronto, December 2006), the Internet circumvention software. For the exhibition (as well as ongoing analysis in Internet censorship), the Issue Crawler (together with an allied proxy tool we made) was put to a new use: the discovery of previously unknown censored Websites.

The Issue Crawler as it currently looks and feels went online in 2004, and has grown slowly, first working on one and now three servers physically located at the Amsterdam Internet Exchange.

The Issue Crawler currently has over 1,000 users world-wide, with the majority from North America and Europe (see list). It is user-supported software, where the support is drawn mainly from institutions - universities, foundations and NGOs, some 25 in total (see list). It is gratis and open to the public upon registration. Public relations firms are turned away, unless for personal use. After login the Issue Crawler map archive (with some 5000 maps) is also open to all users, searchable by map name or URL on map. Individual users cannot be profiled by searching.

ISSUECRAWLER SUPPORTERS (2002-2007)

Soros Internet Program, New York (seed grant)
Open Society Institute, Budapest/London
Ford Foundation, New York
MacArthur Foundation, Chicago
University of Pennsylvania (Annenberg School)
Ecole des Mines, Paris
University of Washington, Seattle (Political Science and Communication Studies)
University of Amsterdam (Philosophy, Media Studies, and the International School)
University of Cardiff, Wales (Economic and Social Research Council)
Cardiff Law School, Wales
University of Alberta, Canada (Political Science and Humanities Computing)
Florida State University, USA (Communication Studies)
Michigan State University, USA (Quello Center)
E-Volve Foundation, Philadelphia
Advocacy, Inc., Philadelphia
Center for Arts and Media, Karlsruhe (ZKM)

New Paradigm Learning, Toronto
Netherlands Ministries of Foreign Affairs and of Education, Culture and Science, The Hague
World Wildlife Fund International (WWF), Gland, Switzerland
University of Vienna (Science Studies)
McGill University, Montreal (Social Studies of Medicine)
University of Pittsburgh (International Affairs)
Brunel University (BRESE)
Citizen Lab (University of Toronto)

ISSUE CRAWLER USERS (current) by sld.tld user addresses (<3 not shown)

.com 298	.fr 25	.edu.sg 3
.edu 143	.edu.au 13	.ro 3
.org 99	.co.uk 12	.ac.at 3
.net 56	.es 7	.dk 3
.ac.uk 54	.hu 7	.com.mx 3
.nl 53	.fi 7	.ie 3
.ca 49	.ch 6	.at 3
.it 38	.org.uk 4	.no 3
.de 32	.be 4	