Chapter 6

Mapping public Web space

with the Issuecrawler

6.1. Introduction

Changes in the conceptualization of Web space may be read from the kinds of visualizations made over the past decade. The piece concerns itself with the visualization modules made for one Web 'mapping' device, the Issuecrawler network location software. It briefly periodizes understandings of Web space by examining the contexts in which Issuecrawler mapping modules were conceived and built: the site inlink list (in the « Web as hyperspace » period), the circle map or virtual roundtable (in the period of the « Web as neo-pluralistic space »), the cluster or issue network map (in the « Web as network » period) as well as the geographical map or the distributed geography of an issue (in the current « locative » period). The focus moves from the « metaphysics » of software-made space (sphere, network) to the specific info-political geographies that can be charted with the aid of the Issuecrawler and allied tools.

Chapter written by Richard ROGERS.

6.2. The death of cyberspace

The symbolic end of 'cyberspace' may be located in the Yahoo lawsuit in May 2000, brought before a French court by two French non-governmental organizations, the French Union of Jewish Students and the League Against Racism and Anti-Semitism. The suit ultimately led to the ruling in November 2000 that called for software to block Yahoo's Nazi paraphernalia pages from Web users located in France.

Web software now routinely knows a user's geographical location, and acts upon the knowledge. You are reminded of the geographical awareness of the Web when you type into the browser, google.com, and are redirected to google.fr. Whilst it may be viewed as a practical and commercial effort to connect users with languages and local advertisements, the search engine's IP-geo-location handling also may be viewed as the demise of cyberspace as place-less space [MIL01].[109]. With location-aware Web devices, cyberspace becomes less an experience in displacement than one of re-placement – you are sent home by default.

What might be called « the announced death of cyberspace » or « the revenge of geography » has consequences for any theorising of the history of Web space. The question posed here concerns how Web space is conceptualized by devices that have sought to 'map' the Web, especially without employing conventional politico-geographical cartography or borrowing from geological metaphors, such as thematic islands, peaks or valleys [DOD01]. In the following I treat one device in particular. The Issuecrawler's sense of Web space is explored through a brief history of the visualization modules created for the software – a history that also seeks to periodize understandings of Web space. It does so through a reflection on how the visualizations provided commentary on contemporaneous Web thought.

The Issuecrawler is server-side Web network location software. Input URLs into the Issuecrawler, and the software crawls the URLs, captures page/site outlinks, performs co-link analysis, and outputs the results in lists as well as visualizations. The software was conceived in the mid-1990s at the Department of Science and Technology Dynamics, University of Amsterdam [ROG96], and has a forerunner in the Netlocator, also known as the De-pluralising Engine, built in Maastricht during the Jan van Eyck Design and Media Research Fellowship, 1999-2000.

---

1. Ethnographers of the web have concluded since 2000 that geography is more significant than media. It seems that the local and national dimensions are more powerful than practices specific to media of so-called non localised space.

6.3. Tethering websites in hyperspace through inlinks

The Netlocator began with the insight that Websites (or Webmasters) link selectively as opposed to capriciously. There is a certain optionality in link-making. Making a link to another site, not making a link, or removing a link, may be viewed as acts of association, non-association or disassociation, respectively. Later, we learned through a Georgia Tech study and our own observations and interviews, that hyperlinks are matters of organizational policy, especially for corporations and government [KEH99], [GOV99]. Selective link-making could create space when one conceives of space as that demarcated by limited acts of association. The demarcationist approach performs an important break with cyberspace by suggesting that hyperlinking behaviors dismantle the 'open-ended-ness' of cyberspace, one that informed the idea of 'placeless-ness.'

What types of associations are on display in hyperlinks (‘reading between the links’), and what could be the shapes of spaces demarcated by link associations? In the late 1990s and early 2000s the leading visualizations we discussed were the Plumb Design’s ThinkMap Visual Thesaurus as well as the I/O/D’s WebStalker, followed shortly thereafter by TouchGraph’s Google Browser as well as Theyrule.net by Josh On[110]. All are non-directed graphs, without arrowheads, which is to say that the elements (synonyms, site pages, board members and companies) are associated (and lines are drawn between them), without specifying a uni- or bi-directional association. Undirected graphs, arguably, derive from a path model of the Web, also built into browsers (with the forward and backward arrows), and lead to ideas about every link being a two-way link [NEL99], [BER99]. Seeing the Web in terms of paths is not far-fetched, since one may surf from page to page, and use the browser buttons, or the browser history, to retrace one’s steps and also move ‘forward’ again. However, on the Web, two-way links, it may be observed, are less frequent than one-way links. Viewing any hyperlink as a bi-directional association, we learned at the time, also has its infamous cases, whereby for example a German ministerial site was accused of being linked to a call boy network [MAR00]. The Bundesministerium für Frauen und Jugend linked to a women’s issues info site, and that info site linked to a call boy network. To the *Bild Zeitung* this Web path implicated government.

To stand on the shoulders of Vannevar Bush, Theodore Nelson and other path and hypertext model pioneers would view the Web as pathway space (for the surfer) [BUS45], [MAR99]. The Netlocator (and later the Issuecrawler), however, strove to distance itself from the Web as pathway space, and instead concentrated on the Web as selective associational space (made by Webmaster linking). How does one view

---

2. The *ThinkMap Visual Thesaurus of Plumb Design* is accessible at http://www.thinkmap.com/ and about *Webstalker* of I/O/D at http://bak.spc.org/iod/.

associations? As is well-known, a site's outlinks, most readily in the form of one or more link or resources lists, are viewable to a site visitor. To gain a sense of a site's inlinks, however, requires the use of the advanced search of an engine, or access to the referrer logs of a site. Until the creation of 'trackback,' a feature implemented in the Movable Type blogging software in 2002 that shows backlinks to a posting, inlinks in the late 1990s were not an everyday concern. Only ranking algorithm makers, most notably Google with the PageRank system, made use of them. Nowadays, on the Web as well as in the blogosphere and in online news, devices recommend pages routinely by counting inlinks, e.g., 'most blogged' stories at the *New York Times* and the *Washington Post*. In all, concern with inlinks as a marker of page relevance or reputation marked a major shift in the underpinnings of Web space.

Counting inlinks addressed the site authority problem. Previously, in the mid-nineties the foremost issue concerning search engine developers related to how to separate the 'real name' from the borrowers of the name, e.g., to return Harvard University at the top of the list when Harvard is queried. In leading search engine results (AltaVista's), the "eminent scientist and the isolated crackpot [stood] side by side," as one leading author put it [RHE94]. In their ranking logics, AltaVista granted site owners the authority to describe the content of their sites (in metatags) and their descriptions became the basis for the engine returns. The Web became a space displaying "side-by-side-ness," fitting with contemporaneous ideas about its pluralizing potential [BAR96]. Google, conversely, granted other sites that authority (hyperlinks and link pointer text). Counting inlinks and having other sites grant authority through linking (and naming their links well) form the basis for most search engine algorithms these days, including Yahoo's as well as MSN's. Once a major competitor to automated search engines, the directory has declined.

6.4. The depluralisation of the web

Which links should be counted? The observation we made in the late 1990s was that search engines' 'population' for link-counting was the entire Web (or the percentage of it they were able to index). Instead of focusing on what the 'most influential' [social network metric] are calling the carriers of the term, on the record on the Web (which is how I would summarize the dominant search engine ranking algorithms), we preferred seeking what could be called 'organizational networks.' Insert a set of URLs of organizations working in the same area, and return those organizations (or URLs) which have received at least two links from the starting points.

Thus, like Google's for the entire Web, the Netlocator's (and, later the Issuecrawler's) algorithm for a portion of the Web crucially sought to take into account sites' inlinks. Once the crawling and co-link analytical procedure of the Netlocator completed, a list of sites in the network (the results) were displayed, color-coded as governmental (.gov), commercial (.com), non-governmental (.org), and scientific (.edu), including country-specific sub-level domains. (.Ac.uk, for example, would count as scientific.) When an actor was clicked, the links it *received* were highlighted. It was not just inlink counts, but types of inlinks, that concerned us, however. When showing an actor's inlink types, the Web could be made into an actor reputational space by showing which links a site received.

How and why do sites link [PAR05]? The Netlocator-related link language for outlinks and inlinks provided a schematic for linking behavior generally, according to domains as well as further qualitative characterisations (see Figure 6.1). In one of our first extended case studies, on genetically modified food, inlinks and outlinks provided actor profiles according to types of links received and given. For example, three corporate sites were compared; the sites' respective standings differ according to the types of links received, and sites' respective display of awareness according to types of links given. One corporation has a different standing by virtue of receiving links from NGOs and government, as opposed to from other corporations only.

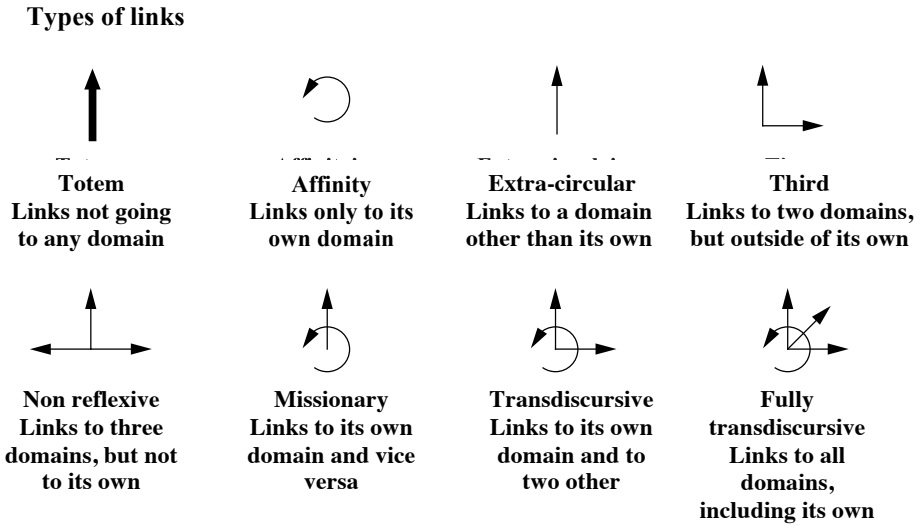

Figure 6.1. *Hypertext link language of actors, Govcom.org*, *Design and Media Research Fellowship, Jan van Eyck Academy, Maastricht, 1999*

Links are also classified qualitatively: cordial, critical or aspirational. Cordial links are the most common – to project partners and affiliates and other friendly or respected information sources. Critical links, largely an NGO undertaking, have faded in practice, and aspirational links are made normally by smaller organizations to establishment actors, often by organizations desiring funding or affiliation. In the case study on genetically modified food products, the « inspirational » links are as follows: Novartis provides a link to Greenpeace. Greenpeace does not provide a link back. Greenpeace and Novartis provide a link to the government. The government does not provide a link back.

In 1998, before the existence of Netlocator, the shape given to the visualization of associational linking in Web space was based, initially, on astronomical charts. Generally, thinking in terms of the Web as a universe (to be charted) coincided with early ideas of the Web as a hyperspace, where one would jump from one site to another at a great distance. Google's "I'm feeling lucky" button also played upon the trope of hyperspace and the famed hyperspace button (from the Asteroids arcade game by Atari, released in 1999). In the period of starry night site backdrops and random site generators, Web sites, arguably, appeared untethered, individual stars whose relationships could be charted (and constellations or configurations perhaps named).

The circular maps that we have created from data output by Netlocator (lists indicating links between pages and sites) also evoked the 'sphere' of public sphere theory The idea that the Web was or could be made into a pluralizing space, where familiar hierarchies of credibility may be challenged, became the focus of our visualization work. With the GM Food (1999) and Russian HIV-AIDS maps (2000), we sought to show interlinkings between sites in a kind of virtual roundtable [ROG00]. What if the Web, according to network inlink counts, were to determine who would sit at the table, instead of more familiar agenda-setters? Significantly, however, not all the actors had the same standing at the table – some receive more links than others and thus grow larger in size. The links between the actors are considered to be entanglements. Are the linked actors all on the same side? Would only the largest nodes speak, and the smaller ones keep still? Thus our roundtable was not flat; it had complications, which we sought to capture in the notion of the 'De-pluralizing Engine,' the other name for the Netlocator. In fact, the Web should not be seen as a pluralizing space by itself, for it is creating hierarchies through inlink counts generally, and through inlink counts from most influential actors (the basis of Google's PageRank).

In the circle maps, especially those auto-generated in what came to be known as the vanilla version of the Issuecrawler (2001), the de-pluralizing spirit continued with built-in notions of a core network and a periphery, where the latter, called the 'waiting room,' comprises those actors (or sites/pages) not quite receiving enough links to sit at the table. A variation on aspirational linking was in evidence in the visualization, as it showed only links from periphery to core, and not from core to periphery. Thus the peripheral link showed a desired belonging to the core, as of yet unachieved owing to lack of sufficient inlinks. Referring to the mapping practice, Noortje Marres prefaces her PhD dissertation with the following remark: "When we [took] to the Web to study public debates on controversial science and technology, we [found] issue networks instead" [MAR05]. Notions of the Web as debate space, with the virtual roundtable construct (however much we strove to complicate it), did not fit with the empirical findings. Even when we endeavored to *make* the Web into a debate space, by harvesting text from organizations' specific, issue-related deep pages, we found only statement juxtapositions – comments by organizations on a particular statement, but scant inter-organizational exchange. Organizations would release views on an issue on their Websites (which we would capture), but forums and other dialogue spaces were not used by what could be construed as the parties to a debate. The Web could not stand in for a building – or an event where debating parties could gather. (Certain authors also began to discuss our work as evidence that the Web (or Net) should not be construed as a public sphere [DEA02].) The Web as neo-pluralistic space had come to an end.

6.5. The web as (issue) network space

Not only with the circle maps but also with the cluster mapping module made by Andrei Mogoutov[111], the final version of Issuecrawler (2003), with its instructions of use, is to be described as an 'issue network' location and visualization machine. The understanding of the web as a network space, as opposed to a virtual space or online community space, is initially linked to the distinction between a multi-site analysis and a monosite analysis. (« Online communities » these days are still generally geographically concentrated and located on one site: for example, Hyves, in the Netherlands, Facebook in the United States, Orkut in Brazil, Cyworld in South Korea, and Lunarstorm in Sweden.). When performing multiple site analysis, with the Issuecrawler, the crawling and co-link analysis return the sociable and the under-socialized, so to speak, in the same space. (Thus we achieved a new form of 'side-by-side-ness.') In terms of types of associations (found in Web space and network mapping more specifically), issue networks may be distinguished from popular understandings of networks, and social networking, in that the individuals or organizations in the network neither need be on the same side of an issue, nor be acquainted with each other (or desire acquaintance) [MAR06]. Actors may be antagonistic, oppositional, adversarial, unfriendly, estranged. Additionally, unlike social networks, issue networks do not privilege individuals and groups, as the networks also may be made up of a news story, a document, a leak, a database, an image or other such items, found on individual pages of Websites. (Thus the Issuecrawler considers 'deep pages' as significant for the study of issue networks.) Taken together these actors and 'argument objects' serve as a means to interrogate the state of an issue either in snapshots or over time. 'Issue states' may be gauged, initially, by taking note of the network's actor composition [ROG04]. For example, we have compared in 2004 queries

3. See Chapter 9.

in Google for « climate change » and « RFID » (radio frequency identification). If we look at the actors who appear most often in the results, for « climate change », we have scientists from the United Nations, governmental agencies, and other actors forming part of the establishment. In the results for RFID, we find the trade press, corporations, lone activists and electronics hackers. If we compare the composition of the network in terms of actors (from the results found by the search engine based on the count of links), it clearly appears that climate change is a more mature issue than RFID.

It is important to emphasize the Web's capacity (with the Issuecrawler) to display configured, professional and publicized culture. The networks or lists that are located rely on public displays of connection (hyperlinks), rather than informal, quiet or old-boy relationships [HOB03]. Indeed, network mapping often has as its goal to make things visible, to reveal non-public relationships, even to dig for dirt. A 2002 search engine query resulted in the newspaper headline: "UN weapons inspector is leader of S&M sex ring" [THE02].) Understandings of the Web as network space, together with the return of the informality of the Web (particular through the blogosphere), have given rise to an investigative outlook. The impulse relates to the Web's street proximity, its closeness to the ground, including the "fact-checking," evidential spirit of the political blogosphere. The Issuecrawler takes into account a sense of a public 'real' – evinced in the making and displaying of a hyperlink. Thus, importantly, the Issuecrawler does not map what is commonly understood as 'virtual space' (as an online game environment). A map of a virtual space would be to a computer or video game what a 'site map' is to a Website, showing the world (or the pages) that have been built and how one may navigate them.

Since 2005 the Issuecrawler has been considered a mapping device for issue professionals and researchers working with that made public. In an effort to make the Issuecrawler's sense of the 'real' even less virtual, the latest visualization module, the Issuegeographer, strives to ground networks (see Figure 6.2). We placed issue network actors on a geographical map to show the proximity (or lack thereof) of the places of actors to the places of issues. The issue of *climate change* has taken place in these places according to the network on the web.

Figure 6.2. *The localisation of an issue. The results of Issuecrawler on climate change prepared by Issuegeographer, 2005*

The focus of the visualization work began to consider actor mobility, whether networked actors move from issue to issue (or whether issues move from network to network). The provocative question read: Do networks form around issues, or are there networks in place that assume issues as they arise? Previously, in social movement research, the idea was mooted that there is 'free-floating movement potential,' in the sense of a given collection of publics which are able to form a movement, with particular conditions. That is, movements are not spontaneous uprisings, but rather more structural phenomena. May the same be said of networks? Are networks simply there, like Websites under construction, waiting for content? Especially global issues may have typical discursive homes, as at (recurring) conferences, summits and other gatherings. Thus, we asked, is there a difference between where is an issue is happening, and where it is currently based? The notion of the 'base of an issue' takes as its point of departure professional circulation, which results in people asking each other, not where you are from, but where you are currently based. With the Issuecrawler in tandem with the Issuegeographer, the Web becomes a space where one can locate where an issue is based [GOV05].

6.6. Conclusion

Issuecrawler visualizations have evolved with conceptualizations of Web space – from hyperspace and cyberspace over public sphere and debate space to network and locative media. In each case the visualizations sought to engage with specific notions of Web space. In the hyperspace period the Netlocator tethered sites by showing inlinks. The Issuecrawler broke with the alleged open-ended-ness of cyberspace by showing how hyperlinks demarcate associational space. It also engaged with public sphere theory (Web as debate space) by unflattening the virtual roundtable, showing over-sized nodes and entangling links. The cluster map module organized actors into a particular kind of network, the issue network, where, with the Issuegeographer, one is able to map the distance between where an issue is happening (e.g., on the ground), and where an issue is currently based (e.g., in a summit network). Recent concrete research projects with the Issuecrawler engage with the current locative media period, where the Web, with tools, may be made to show information politics in specific geographical settings [ROG06][112].

6.7. Bibliography

[BAR 96] BARBROOK R., CAMERON A., « The Californian Ideology », *Science as Culture*, vol. 6, n° 1, p. 44-72, 1996.

[BER 99] BERNERS-LEE T., *Weaving the We : The Past, Present and Future of the World Wide Web by its Inventor*, Orion, London, 1999.

[BUS 45] BUSH V., « As We May Think », *Atlantic Monthly*, p. 101-108, July 1945.

[DEA 02] DEAN J., *Publicity's Secret: How Technoculture Capitalizes on Democracy*, Cornell University Press, Ithaca, New York, 2002.

[DOD 01] DODGE M., KITCHIN R., *Mapping Cyberspace*, Routledge, London, 2001.

[GOV 99] GOVCOM.ORG, The Rogue and the Rogued: Amongst the Web Tacticians, Video, Jan van Eyck Design and Media Research Fellowship, Maastricht, 1999.

---

5. The author wishes to thank Noortje Marres for her critique, as well as Anat Ben-David, Erik Borra, Marieke van Dijk, Koen Martens and Auke Touwslager for their maps and programming. For more information on the software presented here: http://movies.issuecrawler.net/.

[GOV 05] GOVCOM.ORG, The Places of Issues: The Issuecrawler Back-end Movie, in B. Latour and P. Weibel (ed.), *Making Things Public: Atmospheres of Democracy*, Film, ZKM, Karlsruhe, 2005.

[HOB 03] HOBBS R., LOBARDI M., *Global Networks*, Independent Curators International, New York, 2003.

[KEH 99] KEHOE C., PITKOW J., SUTTON K., AGGARWAL G., ROGERS J., *GVU's Tenth World Wide Web User Survey, Graphics Visualization and Usability Center*, College of Computing, Georgia Institute of Technology, Atlanta, 1999.

[MAR 99] MARRES N., ROGERS R., « To Trace or to Rub: Screening the Web Navigation Debate », *Mediamatic*, vol. 9-10, p. 117-120, 1999.

[MAR 00] MARRES N., Somewhere you've got to draw the line: De politiek van selectie op het Web, Masters thesis (MSc.), University of Amsterdam, August 2000.

[MAR 05] MARRES N., No Issue, No Public: Democratic Deficits after the Displacement of Politics, PhD. dissertation, University of Amsterdam, 2005.

[MAR 06] MARRES N., « Net-Work Is Format Work: Issue Networks and the Sites of Civil Society Politics », in J. Dean, J. Asherson, G. Lovink, (ed.), *Reformatting Politics: Networked Communications and Global Civil Society*, Routledge, New York, 2006.

[MIL 01] MILLER D., SLATER D., *The Internet: An Ethnographic Approach*, Berg, Oxford, 2001.

[NEL 99] NELSON T., « Xanalogical Structure, Needed now more than ever: Parallel Documents, Deep Links to Content, Deep Versioning, and Deep Re-use », *ACM Computing Surveys*, vol. 31, issue 4, December 1999.

[PAR 05] PARK H., THEWALL M., « The Network Approach to Web Hyperlink Research and its Utility for Science Communication », in C. Hine (ed.), *Virtual Methods: Issues in Social Research on the Internet*, Berg, Oxford, p. 171-181, 2005.

[RHE 94] RHEINGOLD H., *The Millennium Whole Earth Catalog*, Harper, San Francisco, 1994.

[ROG 96] ROGERS R., « The Future of Science & Technology Studies on the Web », *EASST Review*, vol. 15, n° 2, p. 25-27, June 1996.

[ROG 00] ROGERS R. (ED.), *Preferred Placement: Knowledge Politics on the Web*, Jan van Eyck, Maastricht, 2000.

[ROG 04] ROGERS R., « The Departure of Science from the RFID Issue Space: Mapping the State of an Emerging Sociotechnological Controversy with the Web », *presentation done at the EASST Conference*, Paris, 28 August 2004.

[ROG 06] ROGERS R., Mapping Web Space with the Issuecrawler, available at: http://www.govcom.org, 2006.

[THE 02] THE DAILY TELEGRAPH, 30 November 2002.