

"The BlogosphereMap" [pre-final Version]

Gernot Tscherteu,
realitylab / Austria,
Hollandstrasse 8/2, 1020 Wien
gt@realitylab.at,
www.realitylab.at,
+43 1 218 63 00 99

Christian Langreiter,
synerge / Austria,
6336 Langkampfen
chris@langreiter.com,
www.langreiter.com,
+43 5332 88 049 77

Abstract:

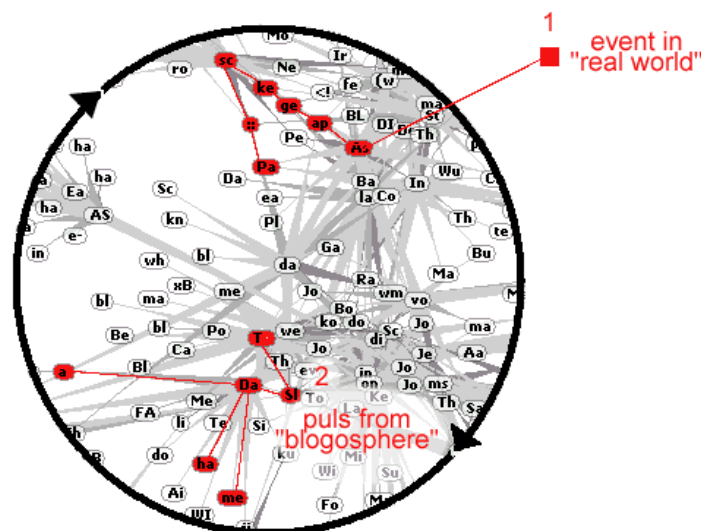
The Internet like any other media is not a mere reflection of reality but a reality on its own: It represents a new cultural sphere and has its own characteristics with respect to communicative codes and knowledge production. The phenomenon of weblogs show that media are not necessarily pointing to "everyday reality" and events in "real world"; weblogs exhibit a remarkably self-referential quality as well: They regularly point to other weblogs where an event has been already intellectually processed, and the way it has been processed - its reception - is to certain extent as or even more important than the original event itself. This process of self-reference can be found in any vivid system and it is believed to be closely related to the emergence of consciousness and intelligence. The interconnectivity of weblogs reminds us the way our neuronal system works. If we look at tools like Blogstreet (in connection with TouchGraph based visualization) and expand several nodes in the neighbourhood of a weblog the apparent similarity is amazing: We find a network of "processing units". Our concept of the BlogosphereMap develops this approach a bit further: What if we could see not only a static map but also (quasi-neuronal) activities on it - the spread of news and thoughts within the neighbourhood of a weblog.

the question whether we accept the underlying theoretical concept (which hardly anyone knows anyway) but more with the fact that it evokes a generally approved perception of networked structures. Finally we will work with visualisations and programmes that have become conventions for whatever reason. It is evident that these conventions will shape our notion of the internet and the blogosphere - and by participating in the invention of new forms of visualisation one participates in the construction of a new reality.

2. Order as Representation vs. Order as Construction

The way we represent and order information reflects our perception of reality. In most cases "reality" is treated as something that exists independently from us and that can be mapped into models by appropriate means of observation and measurement. Others, commonly labelled as constructivists¹, believe that reality is a construction produced by means of our physical and cultural perception. Accordingly for most people "order" should accurately map reality whereas constructivists consider "order" more a tool for the comprehension and manipulation of reality.

Weblogs is probably the best media to demonstrate these two functions of order: On the one side a weblog represents a protocol of an author's experience within and without the internet, but on the other side these observations itself become part of the media and influence its evolution. This self-reference is fascinating and it attracts our attention: Within the movement of Weblogs a new kind of metamedia is arising. It perceives and processes the growth of the net as well as events outside of it.



This self-referring system is neither built up by single individuals nor is it centrally coordinated. It happens in an uncontrolled, self-organising way and it results in spontaneous order. Weblogs create a sphere of recursion and interconnectivity, that is more dense and dynamic than the rest of the Internet, like the number of neurons and its interconnexions are higher and denser in the brain than anywhere else in the human body.

¹ introductory constructivist literature:

Maturana, R. Humberto; Varela, Francisco J. (1987): Der Baum der Erkenntnis. Die biologischen Wurzeln des menschlichen Erkennens. Bern: Scherz.

Varela, Francisco (1981): Der kreative Zirkel - Skizzen zur Naturgeschichte der Rückbezüglichkeit: In: Watzlawick (Hg.), Die erfundene Wirklichkeit. München: Piper, 1981.

von Foerster, Heinz (1985): Sicht und Einsicht. Braunschweig/Wiesbaden: Vieweg.

I developed a constructivist media theory (german only):

Tscherteu, Gernot (1998). Medien.Kulturen.Wirklichkeiten, zur kulturellen Bewältigung alter und neuer Medien. Dissertation: Universität Wien. (<http://www.realitylab.at/dizz.shtml>)

The blogosphere - like our nervous system - is not only processing external stimuli (1) but also pulses from within the nervous system itself (2): for example sensations, thoughts, and dreams. Correspondingly the blogosphere is not only processing news from the real world but also already processed news and "processings of processings" culminating in ideas and concepts, that were created in the blogosphere itself. This self-reference is the main reason for the attraction of weblogs. It makes us act as individuals and as social beings unlike machines that are reacting on external stimuli in fully determined and predictable patterns. In contrast to any automated filter and any artificial system of order the blogosphere is completely autonomous, i.e. it is driven by its own system-immanent rules. This autonomy is one of the biggest treasures of the weblog movement, because it means that webloggers themselves control the evolution of the blogosphere.

3. Planned vs. self-organising knowledge-management.

The blogosphere is constantly processing information. As a macro-effect it produces a system of order which easily overcomes difficulties that would necessarily occur in any planned system, like e.g. TopicMaps:

"Topic maps are a new ISO standard for describing knowledge structures and associating them with information resources. As such they constitute an enabling technology for knowledge management. Dubbed "the GPS of the information universe", topic maps are also destined to provide powerful new ways of navigating large and interconnected corpora"
(Pepper, <http://www.ontopia.net/topicmaps/materials/tao.html>)

The screenshot shows the 'omnigator VI' interface, powered by the 'ontopia topic map engine'. The main content area displays information for 'Puccini, Giacomo', identified as a 'composer'. The interface is divided into several sections:

- Names:**
 - Puccini, Giacomo
 - Giacomo Puccini - Scope: *normal form*
 - Puccini - Scope: *short name*
- Metadata:**
 - born:** 1858 (22 Dec)
 - died:** 1924 (29 Nov)
- Related subjects:**
 - born in:** Lucca
 - composed:** Edgar, Gianni Schicchi, Il Tabarro, Il Trittico, La Bohème (Puccini), La fanciulla del West, La rondine, Le Villi, Madame Butterfly, Manon Lescaut, Suor Angelica, Tosca, Turandot
 - died in:** Brussels
 - exponent of:** verismo
 - pupil of:** Ponchielli, Amilcar
- External resources:**
 - article:**
 - file:/C:/ontopia/topicmaps/opera/occurs/snl/puccini.htm - Scope: *offline; Store Norske Leksikon*
 - <http://www.ontopia.net/topicmaps/examples/opera/occurs/snl/puccini.htm> - Scope: *Norwegian; online; Store Norske Leksikon*
 - gallery:**
 - file:/C:/ontopia/topicmaps/opera/occurs/puccini-gallery.htm - Scope: *offline*
 - home page:**
 - file:/C:/ontopia/topicmaps/opera/occurs/hnh-puccini.htm - Scope: *Naxos; offline*
 - <http://www.hnh.com/composer/puccini.htm> - Scope: *Naxos; online*
 - <http://www.r-ds.com/opera/pucciniana/gallery.htm> - Scope: *online; OperaResource*
 - illustration:**
 - file:/C:/oks-professional/jakarta-tomcat/webapps/omnigator/WEB-INF/topicmaps/occurs/composer/puccini.gif
 - sound clip:**
 - <http://www.puccini.it/files/vocepucc.wav> - Scope: *Centro studi Giacomo Puccini; Italian; online*
 - web site:**
 - <http://www.puccini.it> - Scope: *Centro studi Giacomo Puccini; Italian; online*

At first sight everybody would agree that a unified large-scale encyclopedia of all web content would be a great thing, but having a closer look at it such a system is not very likely to come about because of the following reasons:

- A central coordination is required at least in respect to standards and formats

- The definition of order structures requires domain experts
- one can doubt that a single "objective" order will be accepted by everyone
- probably the most striking argument is that topics - especially in the internet - are very time-sensitive. The connections between topics vary over time.

In contrast to TopicMaps weblogs work as a distributed filter that is constantly processing news. It works because thousands of individuals participate in this process and abide a handful of conventions². These two factors (the big number of users and the conventions) are the key to the weblog movement's success, by generating a self-organising knowledge-management. This system needs not to be created and standardised, it is already there and it is self-evolving.

² commonly accepted conventions and standards:

1. Entries are sorted by date; Latest entries are displayed at the homepage.
2. Each entry has a unique URL (Permalink)
3. older entries can be found in the archive
4. a blogroll lists weblogs and sites frequently visited by the author
5. rss feeds simplify news aggregation.

4. Tools for Comprehending the Internet

As the production and processing of news takes places in a distributed way a number of indices and methods to track activities in the blogosphere and to filter interesting weblogs and news has been created.

a. Blogwise (<http://www.blogwise.com/>)

Firstly this done with registers, like blogwise, which arrange weblogs by different criteria like e.g. topic or country of origin.



blogwise 

3403 sites listed. (155 in last 7 days) [Add Yours Now!](#)

Home
Top Blogs
Search
Submit Your Blog
List by Country
List by Keyword
FAQs
About

Blogwise is...

...a new site bringing you a collection of categorised blogs from around the world. Whatever your taste, interest and culture - there's bound to be a blog out there written by a person just like you. If not, then why not start one yourself!

May 1, 2003

If you have your location listed in your blog by way of ICBM headers (as used by [GeoURL](#)), Blogwise will now show your blog's registered location. I have added (I think) practically every major place in the world.... apart from the USA and Antarctica (don't laugh, they're on a seperate database).

'Location' is now shown in the Blog Info page, which can be gotten to if you click the little question mark that appears next to each blog. For help on adding latitude and longitude to your blog please read [GeoURL's submission page](#) (it's a good thing to add your site to their database as well). For more information on how it is used here, please read my [FAQs](#).

Incidentally, of all the blogs on the database I count 18 that have locations listed, come on guys! - I'll let you know how many (if any!) that jumps by by next week.

[More from the Blogwise Blog \(English language only\)](#)

Recent Additions

Random Blog

[Unknown Rhythm](#)
Blog of many flavours: Tarot Writing Pets 
Country: United Kingdom

[Refresh the page for another random blog](#)

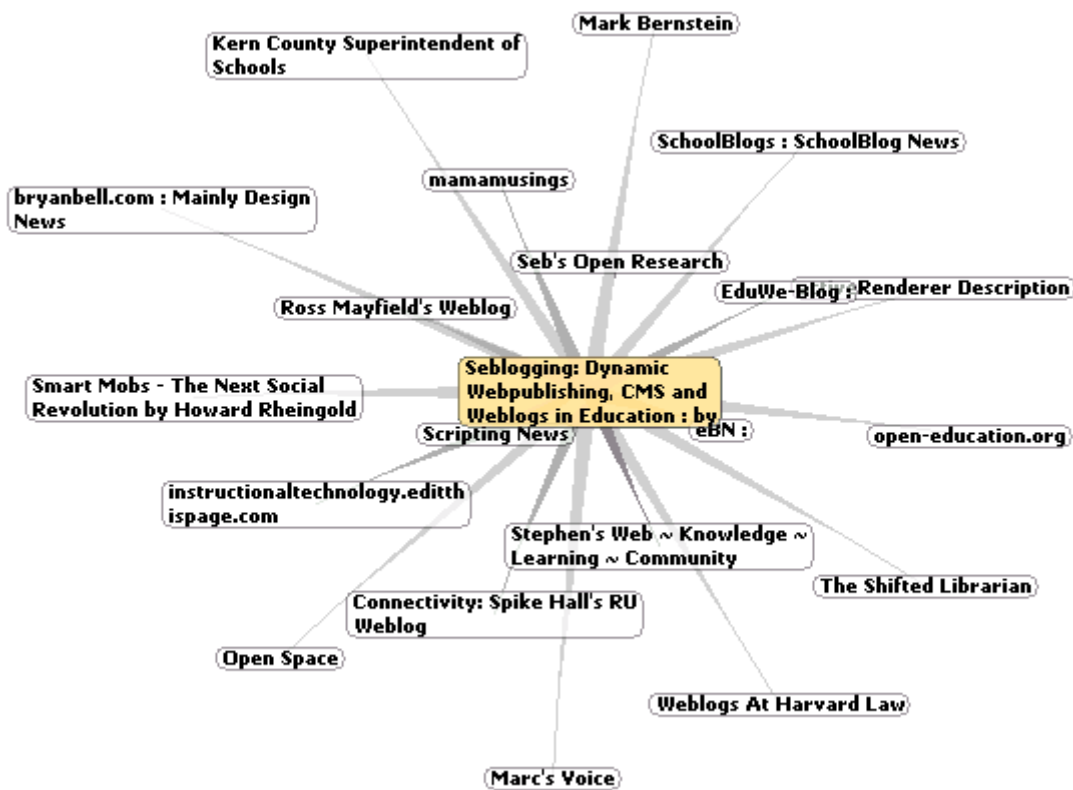
Search

Blogs by Country	Blogs by Keyword
United States (1491)	music (536)
Brazil (404)	news (361)
United Kingdom (312)	blog (331)
Canada (205)	politics (323)
Spain (100)	life (308)
Finland (68)	computers (233)
Iran (Islamic Rep.	humor (188)
	personal (174)

This form of indexing is driven by a rather classical approach of ordering that reminds us of TopicMaps, with all its advantages and disadvantages. Possibly the most striking disadvantage: Like other indices blogwise is not time-sensitive: It is not apparent what the currently active blogs are or which one are neglected at present.

b. **Blogstreet** (<http://www.blogstreet.com/>)

Blogstreet follows a completely different approach, that is based on the link structure of weblogs. The following screenshot maps the neighbourhood of a weblog, which - in this case - is defined by the blogroll of a weblog.³



Blogstreet is the perfect tool for exploring your own neighbourhood, especially neighbours of neighbours - because each node is expandable. Thus you get a brilliant overview of who is reading whom and who is quoting whom. Blogstreet gives us a impression about the spread of news though it doesn't inform us about single news-items.

And that's exactly what Daypop does. (see next page)

³ "we take into account *whom you blogroll* and *who blogrolls you*, add to that a few other factors, give weightage to all the factors and there it is - the Neighbourhood." (<http://www.blogstreet.com/faq.html>)

Daypop is crawling about 17000 websites on a daily basis, some of them even more often.⁴

"Word Bursts" are heightened usage of certain words in weblogs within the last couple days. They are indicators of what webloggers are *writing* about right now, in contrast to [Top 40](#) and [Top News](#) which are indicators of what webloggers are *linking to* currently. Word Bursts can frequently indicate current events of interest that are not usually accompanied by links. Sample weblog posts accompany each word burst⁵

d. [Blogdex](http://blogdex.net/) (<http://blogdex.net/>)

Blogdex offers a service that at first sight seems to be similar to daypop: It indicates the news-items that have been linked most frequently. Interestingly the results can be quite different to those of daypop. The reason is probably to be found in the different data pool (other weblogs and sites) of the two services.⁶

THE WEBLOG DIFFUSION INDEX - [HTTP://BLOGDEX.MEDIA.MIT.EDU](http://blogdex.media.mit.edu)

blogdex

Welcome to the newly reconfigured Blogdex. The frontend has changed dramatically, as well as pretty much everything under the hood. Please excuse the bugs while I get things up to speed.

1. **Dante's Inferno Test - Impurity, Sin, and Damnation**
4degrez.com/misc/dante-inferno-test.html
 >> [track this site](#) | 27 links
2. **Loyalty Day, 2003**
[whitehouse.gov/news/releases/2003/04/20030430-26.ht...](http://whitehouse.gov/news/releases/2003/04/20030430-26.html)
 >> [track this site](#) | 24 links
3. **AlterNet: Patriot Raid**
alternet.org/story.html?StoryID=15770
 >> [track this site](#) | 23 links
4. **BBC NEWS | Health | Boy 'pregnant' with twin brother**
news.bbc.co.uk/2/hi/health/2987963.stm
 >> [track this site](#) | 19 links
5. **Dante's Inferno Test - Impurity, Sin, and Damnation**
4degrez.com/misc/dante-inferno-information.html#9

INFORMATION

Blogdex News Search
 Add your weblog
 XML: [RSS 0.91](#)
 Contact Blogdex

A YEAR AGO TODAY

1. **Site Barks About Deep Link**
2. **Jenin 'massacre...**
3. **Jonah Goldberg's Gold...**
4. **CNN.com - Girls made...**
5. **Save Internet Radio...**
6. **Google to Power Searc...**
7. **Things my girlfriend...**
8. **Barking Dogs - Taking...**
9. **What Blogging Archety...**
10. **Provocative exchange...**

more history coming soon!

⁴ "Weblogs are generally crawled every twelve hours unless they use weblogs.com's ping notification. When a blog "pings" weblogs.com, it's announcing that there have been updates to the blog. Weblogs.com then publishes this information and Daypop makes use of it every hour. This means reduced latency between the time when a blogger posts and the time when Daypop can return that post in its search results. RSS feeds and RSS feed update information are provided by NewsIsFree which regularly crawls existing feeds and also constructs (scrapes) some based on HTML site content. NewsIsFree publishes this update information and Daypop makes use of it every hour to crawl RSS feeds. Communication between sites using XML/RSS is the first step towards a real-time search engine. "

further infos about daypop search: <http://www.daypop.com/info/technology.htm>

⁵ <http://www.daypop.com/burst/>

⁶ Unfortunately there is no aggregated technical documentation about blogdex.

But there is an other quite fascinating feature of blogdex that seems to be quite unique: It indicates the spread of specific news items over time: When does a specific news item or posting appear in other weblogs? It doesn't necessarily mean that you can track who is quoting whom - but nevertheless it allows you to draw some conclusions about the spread of news.

THE WEBLOG DIFFUSION INDEX - HTTP://BLOGDEX.MEDIA.MIT.EDU

blogdex

4degreez.com/misc/dante-inferno-test.html

This site was found on following weblogs, on the given date:

May 02, 2003

kungfukitten.diaryland.com

Dante's Inferno Hell Test

[» track this weblog](#)

4:06 AM

jmtee.blogspot.com

Dante's Inferno Hell Test

[» track this weblog](#)

4:02 AM

michaeljennings.blogspot.com

Dante's Inferno Hell Test

[» track this weblog](#)

3:55 AM

thundersnake.com/news.htm

The Dante's Inferno Test

[» track this weblog](#)

3:45 AM

blog.onemonkey.org

Dante's Inferno Test

[» track this weblog](#)

3:44 AM

MORE

[Recent Citations](#)
[Site History](#)

INFORMATION

[Blogdex News](#)
[Search](#)
[Add your weblog](#)
XML: [RSS 0.91](#)
[Contact Blogdex](#)

5. BlogosphereMap

In our design for the BlogosphereMap we have tried to combine the aesthetics and the functionality of the tools presented above and at the same time to go a step further.

a. Objective

The objective of the BlogosphereMap is to map the spread of a specific news-item (or the underlying thought) over time. While Blogdex delivers a chronological list that potentially covers all weblogs in the data pool, our BlogosphereMap concentrates on the spread of news items in a specific neighbourhood. Similarly to Blogstreet this neighbourhood can be explored interactively with a graph based interface. Like in a neuronal net weblogs flash (fire) when a news item has been posted.

A news-item can not only characterised by a link, but by a sequence of words as well; i.e. it is possible to track the origin and the spread of quotations - extensively used within the Radio Userland community - through it doesn't contain a URL.

b. Interface

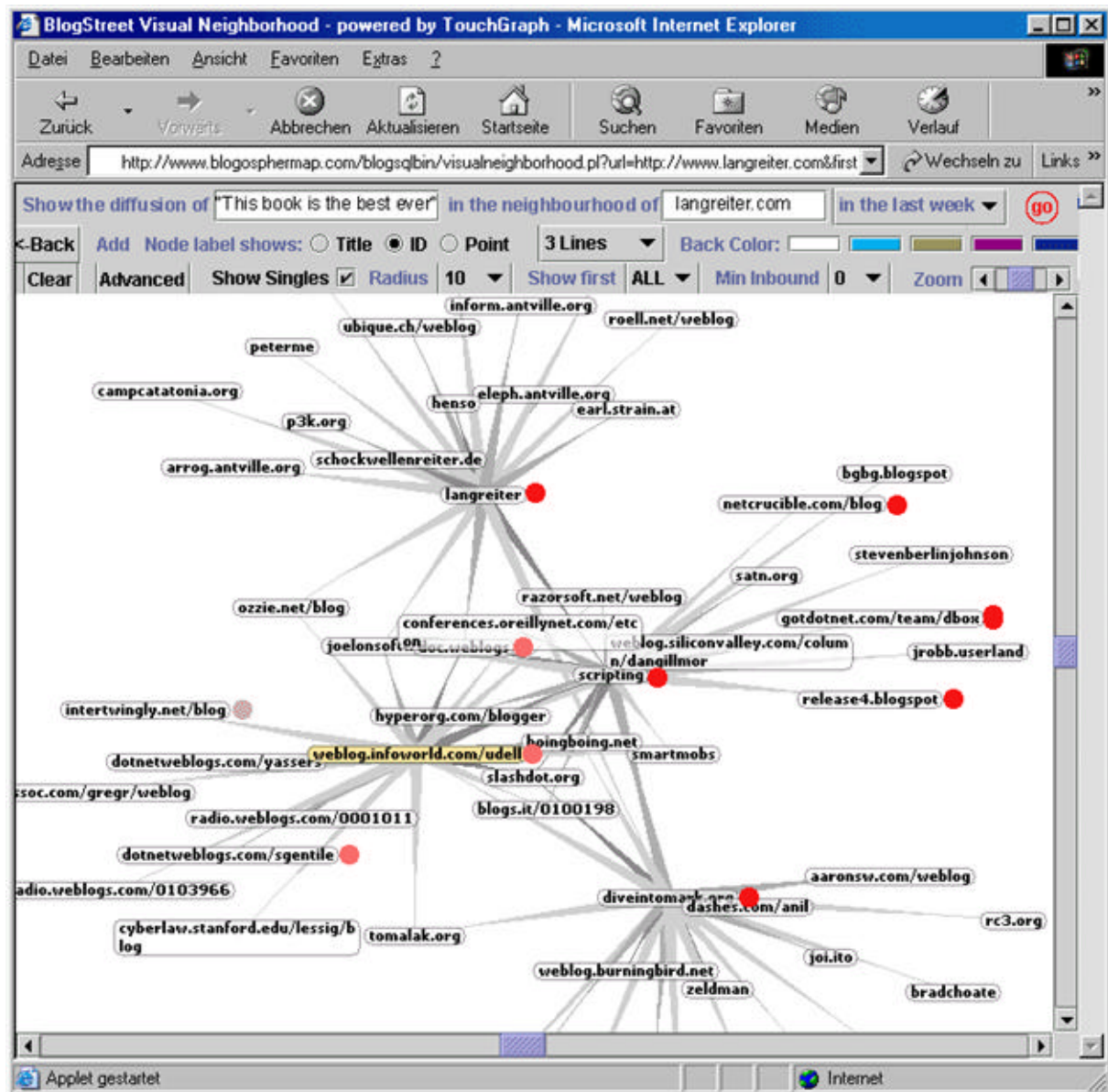
A search requires two parameters.

- a sequence of words
- an interval of time

The search result is mapped in form of an animation that shows the spread of the searched news-item within the chosen time interval. A flashing spot (=weblog) indicates, that a news item has been published at the indicated time. The spot glows on for a certain time in order to let the user recognise a pattern of dissemination.

If spots are flashing one after another it doesn't necessarily mean that the second is quoting the first one, because the system is only able to track a sequence in time. But the grouping of weblogs within a neighbourhood allows such conclusions with a relatively high probability.

(Screenshot see next page)



c. technical realisation

A grabber is collecting about 5000 weblogs and rss-feeds several times a day. They form a data pool for the search, that returns for each weblog the exact time when the searched sequence of words has been published. This result is returned the form of a chronological list showing the sequence of publishing of the searched sequence. At the same time the neighbourhood of the searched weblog is being mapped. In the resulting graph the diffusion of the searched news item is being displayed.

6. Conclusion

All the tools cited above have in common that they are not simple illustrations of information flow but tools as well; tools that allow us to sensually explore a space of information. Its most important objective is to sensually experience the blogosphere within the narrow boundaries of a conventional PC-based interface. Considering these limitations the well-known Touchgraph-based applications do quite a good job in setting up first experiences in an intrinsically invisible space. The haptic sensation of a data space creates a definite notion of structure that is beyond direct perception. Such experiences are fundamental for the emergence of "reality" in a virtual space: Sensual perceptions in a recursive self organising network of information may help us to shape up the blogosphere. A truly thrilling reality lab with the potential for extensive cultural consequences.