Buzzaar, An Environment and Tool to Map Your Web

Keywords: web, collaboration, data visualization, mapping, user, knowledge sharing, network

AUTHORS

Katarzyna Boron meetopia- netart Genève cax@buzzaar.net

Alexander Elmer Lausanne a.elmer@buzzaar.net

Adla Isanovic Academy of Fine Art Sarajevo a.isanovic@buzzaar.net

Nathalie Perrin meetopia- netart Lausanne n.perrin@buzzaar.net

Martin Rajman EPFL Lausanne, Switzerland Martin.Rajman@epfl.ch

Tran Huu Duc EPFL Lausanne, Switzerland huuduc.tran@epfl.ch -

ABSTRACT

As Web users, we visit pages and our navigation history is often locally stored in our browser logs thus creating a data set that contains a lot of valuable information, not only for ourselves but also for peer users we might be willing to share this information with.
How can we best exploit this information? How can we give it a meaning and transform it into knowledge that can be shared with others?
Buzzaar is a tool and a web site that provide answers to such questions. It is free and available for all.
Buzzaar users can interact with their browsing history and transform it into meaningful, dynamic webmaps that intuitively and dynamically represent their interests and navigation habits.
Buzzaar webmap is an efficient representation of web pages visited by a user. It can be easily modified and tailored to specific user needs. In particular, users can tag, filter, sort the visualized data and customize its appearance.
These webmaps can also be merged or projected into each other to identify commonalities or differences, thus allowing their owners to better understand and exploit the knowledge they contain.

The Buzzaar website also allows anyone to register to the Buzzaar network, within which peers can share their webmaps and collaboratively build a common webmap representative of their community as a whole. This common webmap provides a useful context to better interpret each peer's individual webmap.

Info: www.buzzaar.net

INTRODUCTION

The development of digital technologies has brought fundamental changes into our lives, culture and society. Social networking media (i.e. Twitter, facebook), user generated sites (i.e blogs, Wikis, YouTube, Flickr, delicious,etc.) and peer-to peer networks have contributed to the development of a new participative culture where users have access to tools to create and share contents with peers.

As web users, we spend significant time online, searching for contents and visiting numerous web pages. These actions leave traces that can be collected, typically within web server logs. These data represent potential sources of valuable information for various private and public purposes.

Commercial companies (e.g search sites such as Google, Yahoo!, sites such as Amazon or specific purposes sites such as Compare.com) have well understood how useful the above mentioned data can be if properly analyzed and transformed into exploitable information about their targeted audiences. However, in most cases, users cannot access the generated information that is kept private although the original data have been created by the users themselves.

On the other hand, a lot of personal navigation data is locally stored in users' browser logs (navigation history) that contain visit frequencies and, if specific browser plugins are used (voyage and buzzaar), visit co-frequencies. However, there are currently no popular and easy to use tools to exploit these data and they remain quite difficult to share in an efficient way (sites such as Delicious do allow to share bookmarks but do not provide any intuitive graphical visualization nor they provide mechanisms for users data aggregation).

We strongly believe that data created by users should return to users who should be able to interact with it and share it with others, in order to build a culture of commons;[1] a participative and distributed knowledge, part of our digital culture. Furthermore, there clearly is a need for web activity analysis tools providing new graphical representations which give more sense to navigation data and show relations that are otherwise difficult to perceive. Such tools should allow to create dynamic interactive displays where users can see, explore, sometimes even touch and manipulate contents. By fostering contextualization and conscientiousness, such tools could also contribute to users' awareness and lead to their empowerment.

In this perspective, the goal of the buzzaar project is to build a free and publicly accessible server (*the buzzaar server*) providing:

-a firefox plugin *(the buzzaar toolbar)* allowing users to gather their site visit frequencies and co-frequencies and to store them in a private space on the buzzaar server;

-a java applet (*the webmap manager*), running on the server, allowing registered users to (1) visualize their navigation data in the form of interactive graphical webmaps that intuitively represent their interests and navigation habits; and (2) allow them to share their webmaps and collaboratively build a common webmap

1 **"**The commons is terminology referring to resources that are collectively owned or shared between or among populations." Wikipedia contributors,2010, "The commons" in *Wikipedia, The Free Encyclopedia ,* Available at <http://en.wikipedia.org/wiki/Commons > [Accessed 16.2.2011]

"Culture of the commons" is a concept frequently used by the Creative commons initiative to underline this idea of a public and shared culture belonging to the whole, like a common good. Video *A shared Culture* by Jesse Dylan,at Creative Commons, Available at < http://creativecommons.org/videos/a-shared-culture >[Accessed 16.2.2011]

representing their community as a whole; This common webmap (as any other shared webmap) can then be used by all registered users to be merged or projected into their private webmaps in order to identify commonalities or differences, thus allowing them to better understand and exploit the knowledge they contain. In addition, the common webmap is also accessible to any (not necessarily registered) user through the buzzaar server and as such represents a valuable source of public information, when manipulated with the webmap manager.

The Buzzaar toolbar and the webmap applet are under the GPL license and the content of the buzzaar website is under a creative commons license in order to contribute to the development of a common culture on the web.

The long term goal of the buzzaar projects is to contribute to the development of a widely exploitable set of web data management techniques and tools allowing reliable collaborative creation of publicly shared information, while providing efficient ways to protect users' privacy by keeping their identity and contributions anonymous.

The team of the buzzaar project is composed of researchers from the EPFL Artifical Intelligence Laboratory and the netartists from the meetopia community.[2]

The project is supported by the Swiss Office of Culture, sitemapping project, and EPFL.

The rest of this paper is organized as follows: section 1 presents main concepts (webmaps and webmap manipulation functionalities) developed in the project; section 2 briefly describes the existing experimental buzzaar prototype; section 3 identifies related works; and the last section provides general conclusions.

2 Meetopia is a collectif of european web net artists. Most projects developped on the plateform focus on elements of the net culture as part of a new culture; hybridity, replication, sharing, and reformulated dentities. Internet is understood as a virtual public place for production, broadcasting and sharing. Meetopia.net, Available at <http://meetopia.net/index.php?id=about >

LIA: The Laboratory of Artificial Intelligence of the EPFL (École Polytechnique Fédérale de Lausanne)- is part of the Institute of Core Computing Science in the School of Computer and Communication Sciences at EPFL and focus on algorithms for distributed problem-solving, and methods for recommendation and reputation systems, LIA, *Available at* < http://liawww.epfl.ch/ >

1. CONCEPTS

1.2 Interactive visualization through webmaps

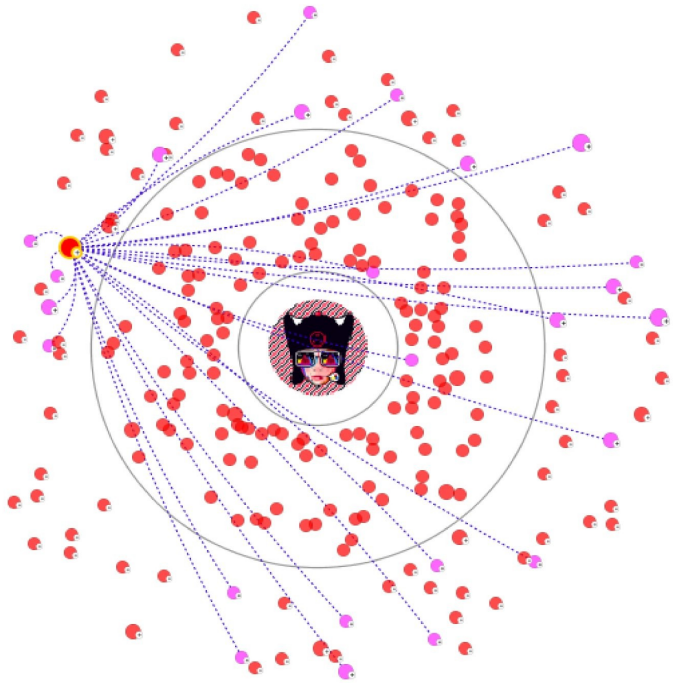


image 1 : Buzzaar webmap

Buzzaar webmaps are visual representations of raw visit frequencies and co-frequencies sent by buzzaar members and stored on the buzzaar server.

As such, webmapping is not new. Several examples of such techniques can already be found (such as TouchGraph, Voyage, etc.; see also the section on related works below). They allow users to efficiently visualize relations between web pages they visit, and in particular, allow to translate huge amount of data into intuitive graphical representations. Such maps can be fully interactive, multimedia, collaborative and sharable, thus enabling users to experiment with their own ways (emotional and intellectual) to knowledge building.

Buzzaar webmaps also correspond to dynamic and interactive web graphs that users can edit, customize, manipulate and share. Edit operations such as filtering and tagging, or specific webmap manipulations such as merge and projection, can be easily performed and provide users with displays of various types of information such as: page popularity (number of visits), co-frequencies (number of times two distinct pages have been both visited in a given interval of time), or categorization (tagging and filtering), and different visual artifacts (size, position, colors and alpha transparency) are used for this purpose.

Therefore, by acting on their webmaps and their different (graphical and textual) components, users can easily see and understand differences among pages, associate meanings to them and better position themselves in their global informational context.

1.2 Webmaps representation and customisation;

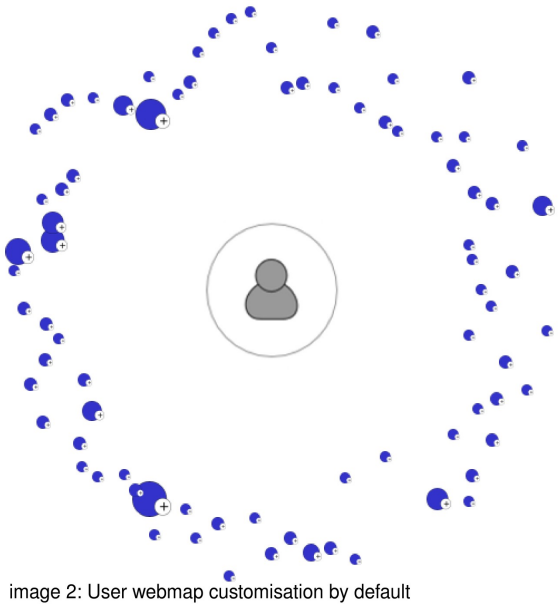
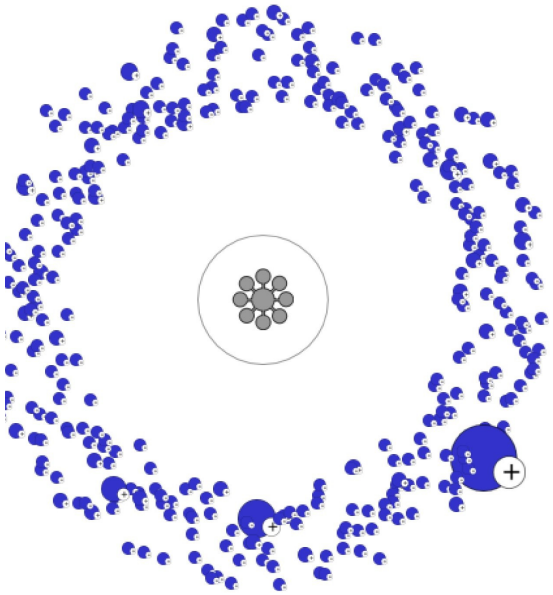


image 2: User webmap customisation by default



mage 3: Common Buzzaar webmap

Buzzaar webmaps are user-centered representations that organize the navigation history of the user around a center that represent the user themselves

The Buzzaar interface is divided into two parts: the list of visited URL (textual part) and a webmap corresponding to a graphical display of URLs. Following standard interface design principles, simultaneity and synchronization are systematically enforced: any change in the textual part (URL list) echoes on the webmap and vice versa.

Further, in both the graphical and textual part, information can be displayed at two different levels of granularity: the URL granularity where the units correspond to individual URLs; and the domain granularity where units correspond to URL domains and aggregate all URLs sharing a given domain. Users can easily switch between two levels of granularity, either by expanding/collapsing an entry in the displayed list or by directly operating on the webmap (image 4a and 4b).

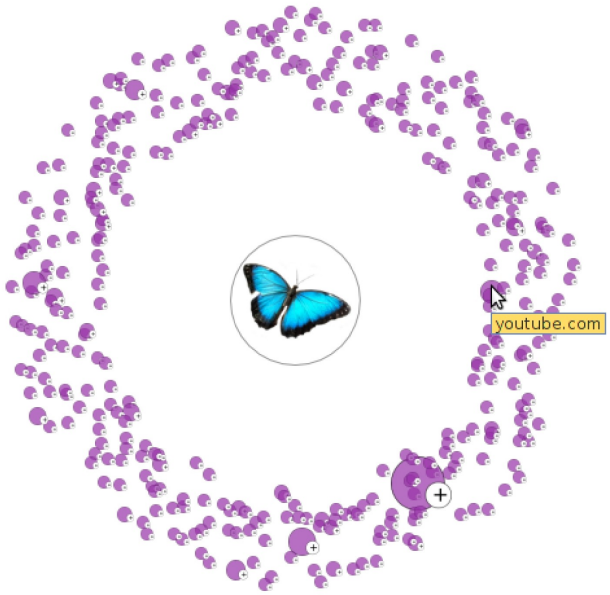


image 4a : Domain mode

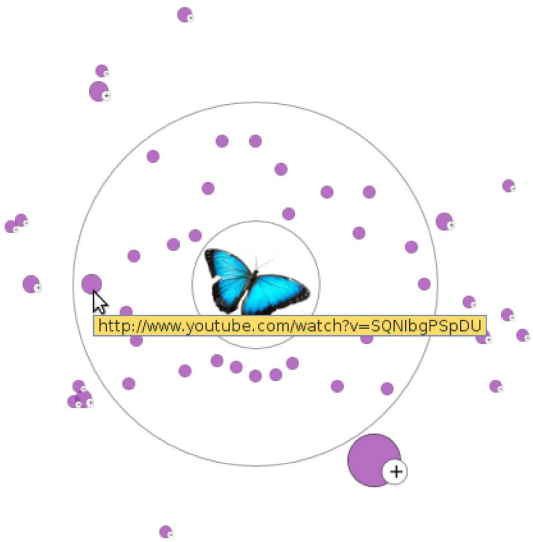


image 4b : Domain mode: URL domain - expanded

In the webmap, URLs and URL domains are represented by differentiated visual artifacts (hereafter called nodes), with a size proportional to the popularity (resp. cumulated popularity) of the associated URL (resp. set of URLs). As this is common in other graphical visualizations (see section on Related works), the representational metaphor stays familiar, intuitive and easily understandable to users.

Finally, webmaps can also visualize undirected (resp. directed) co-occurrence links between URLs, where co-occurrences links are defined as follows: there is an undirected (resp.directed) co-occurrence link between URL A and URL B, if the user has visited URL A and URL B (resp. URL A before URL B) frequently enough, in a given interval of time. In addition, users can also choose to visualize either all co-occurring links (image5) or only those related to a selected subset of URLs and can switch between undirected and directed links (image6). The notion of co-occurring URLs and the visualization of co -occurrence links are important features of the buzzaar webmaps. In particular, they help users to contextualize URLs or URL domains displayed in their webmaps.

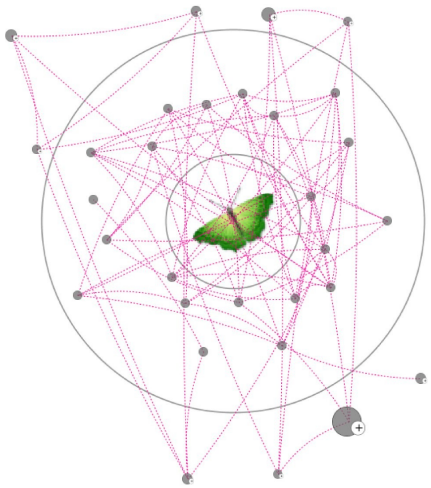


image 5: Co-occurrence: show all

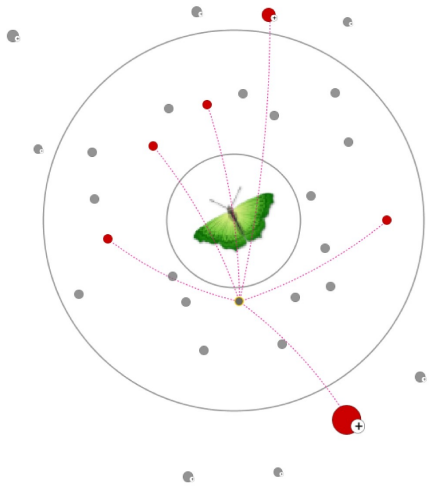


image 6: Co-occurrence: from selected node

In terms of layout, a webmap is divided into four concentric zones, each devoted to a specific type of information:(1) a central zone containing the visual identity of the user (typically, an icon that can be customized); (2) an expansion zone, where aggregating nodes (URLs domains) can be expanded and nodes associated with a given tag can be displayed (see the description of the domain/tag modes below); (3) a default zone, dedicated to the visualization of URL domains (aggregating nodes) or, when a tag or a domain is expanded in the expansion zone, URLs and URL domains co-occurring with the expanded content; and (4) a peripheral zone containing external co-occurring URLs (see the notion of projection in the corresponding section below).

Webmap manipulations

Besides representational aspects described in the previous section, another important feature of webmaps is that they can be associated with various manipulation mechanisms that can be used to better understand and exploit their content. Available manipulation mechanisms can be decomposed into two main types: ones operating on a single webmap (filtering, tagging and customizing), and ones operating on two distinct webmaps (merge and projection).

As the Buzzaar infrastructure allows users to share their webmaps with their peers (typically by sending webmap files over mail), the merge and projection operation are usually performed with one of the webmap being the webmap of the given user and the other corresponding to a webmap that this user has received from some peers.

Filtering Tagging and customizing

Filtering. Webmaps users have the possibility to focus on specific URL name patterns to more efficiently exploit the information in their webmaps. Filtering can be either temporary (typically when exploring a given webmap) or persistent if the user saves the resulting webmap which can then be used independently.

Tagging. When surfing on the web, users often want to categorize information they collect and find important. In this perspective, each web browser provides the possibility to bookmark and tag favorite links, and social networks such as Delicious also offer the possibility to share tags and bookmarks. As a consequence, tagging is nowadays considered as a popular and useful functionality, which allows to give a personal meaning to collected information. Buzzaar therefore provides its users with a tag mode (image 7 and 7b), allowing them to categorize the visualized URLs into topics of interests described by tags that can be further differentiated with colors and shared along with the webmaps users share (typically through a merge operation).

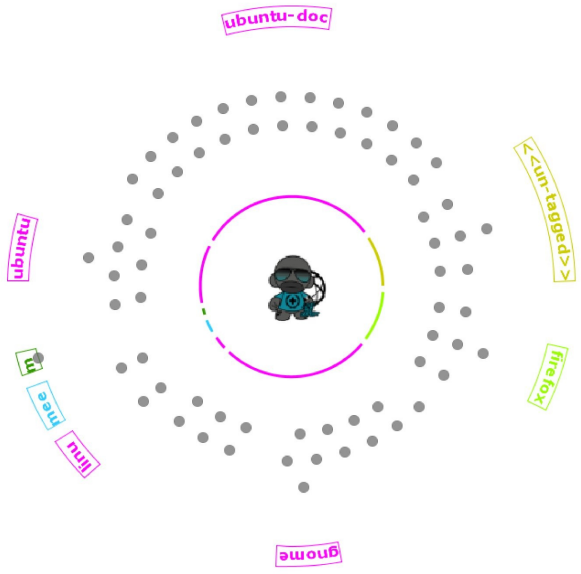


image 7a: Tag mode

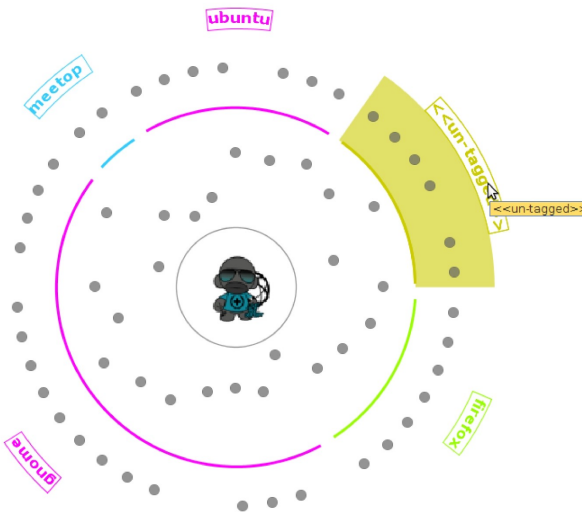
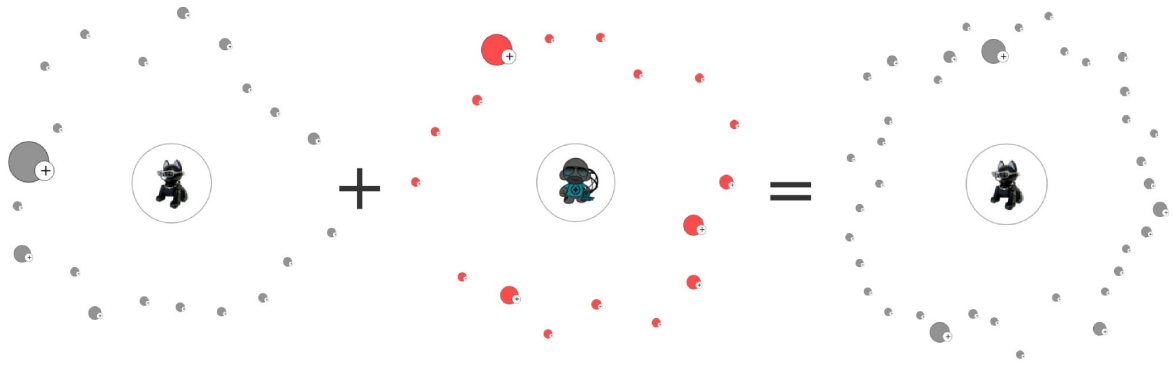


image 7b: Tag expanded

Customizing. Finally, users can personalize the display of their webmaps in different ways. They can change colors of nodes, of co-occurrence links and they can personalize the central icon. Although technically quite simple, this functionality is important, as in social media users are strongly expecting such possibilities. In additions, it makes webmaps more playful and personal, which reinforces users centered nature of the buzzaar approach.

Merge and projection

Merge. Merge allows users to add webmaps (typically ones received from other users) into their own webmaps. In resulting webmaps, URLs, URL frequencies/co-frequencies and tags are cumulated. The merge operation can therefore be used when a group of friends, peers, colleagues, etc. want to put their information in common into a single webmap representing their community as a whole (image 8).



$$x \in X = A \cup B \Leftrightarrow [x \in A \text{ or } x \in B]$$

image 8a, 8b, 8c: Merge operation

Projection.

Projection is a webmap operation that allows users to compare and differentiate information present in two webmaps. In particular, it allows users to easily search for information present in the projected webmap that is relevant to the information present in the webmap into which they are projecting.

More precisely, the projection is a non symmetric operation that can be described as follows: the result of the projection of a webmap A into a webmap B is a webmap C s.t.: the central zone of C is the central zone of A (i.e. the visual identity of the user performing the projection is preserved); the default zone of C contains all the URL/URL domains common to A and B (thus allowing to easily spot commonalities); and the peripheral zone of C contains all the URL/URL domains of B that are co-occurring with some of the nodes in the default zone of C (thus allowing to easily identify information in B not present in A that is potentially the most relevant to A). (image 9)

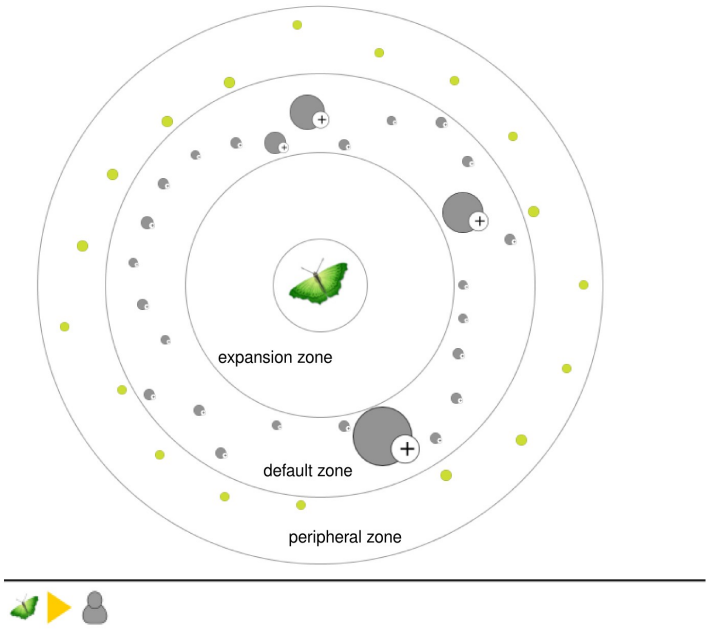


image 9: Projection operation

The projection operation can thus be interpreted as a simple (graphical) equivalent to the recommendation techniques used by many e-commerce companies such as Amazon.

Furthermore, to allow users to more efficiently exploit the potentially relevant information identified by projection, the following visual hints are also provided:

- the external co-occurring nodes are displayed in a specific ***color*** (that can be customized by the user) in order to be easily distinguished from the nodes present in the original webmap; **-** the ***alpha-transparency value*** of all the displayed nodes is proportional to their popularity in the projected webmap; this graphical feature provides users with a very intuitive way to simultaneously assess the importance of displayed nodes in the original webmap (through their size) and in the projected webmap (through their alpha-transparency).

The projection operation does therefore not simply identify which of URLs in the projected webmap are complementary to ones present in the original one, but rather focuses on the complementary URLs that are potentially relevant (because they are co-occurring with one of the URL present in the original webmap), and indicates (through their transparency) how much these URLs are potentially relevant. Similarly, the projection operation does not simply identify commonalities between two processed webmaps, but also indicates how much each of the common URLs displayed in the default zone is important (1) for in the original webmap (through its size); and (2) in the projected webmap (through its transparency). As such, the projection operation represents a very useful tool for users to very efficiently process the information generated by the projection (image 10a and image 10b).

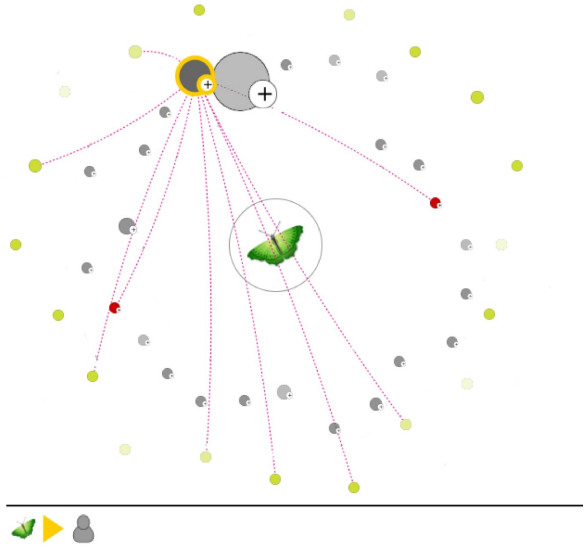


image 10a: Projection in Domain mode

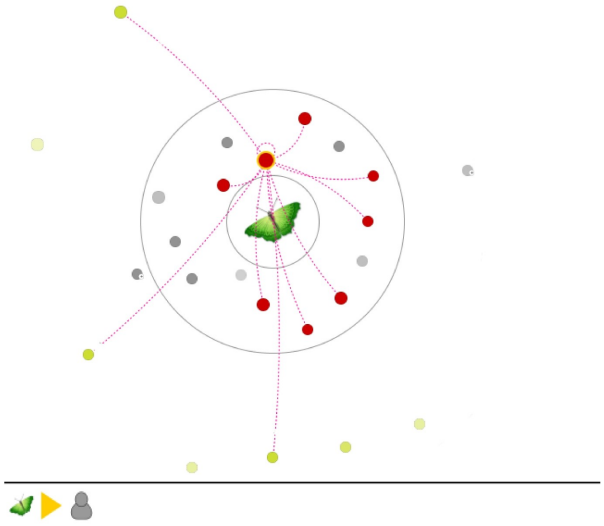


image 10b: URL domain expanded in projection

3. EXPERIMENTAL PROTOTYPE

The Buzzaar project is currently being implemented in the form of an experimental prototype consisting of (1) the Buzzaar server and applet running on a server hosted at EPFL; (2) the Buzzaar toolbar, a firefox plugin allowing users to generate raw visit frequencies and co-frequencies required for them to build their own webmaps; and (3) the Buzzaar website (www.buzzaar.net), maintained by the meetopia community and providing information on the project (up to date news, charter, FAQ, tutorials, forum, webmap gallery, etc.) The Buzzaar toolbar and applet are under the GPL license and the content of the buzzaar website is under a creative commons license in order to contribute to the development of a common culture on the web.

The Buzzaar prototype is freely accessible for anyone, either as guest user or as registered user (member). Guest users can access and explore common webmaps present in the public space of the server. Registered users benefit from an additional private space where they can store and process private webmaps they can produce with the data generated by the Buzzaar toolbar, if installed in their browser.

To prevent the misuse of storage space on the Buzzaar server, registered users can only produce and exploit webmaps generated from web site visit frequency and co-frequency data transmitted to the server by the Buzzaar toolbar (Firefox plugin) when installed in their web browser. The toolbar also provides users with possibilities to filter the data to be transmitted (thus allowing them to prevent the transmission of possibly sensitive information) and the possibility to suspend/resume data transmission at any time. Finally, the data sent to the server by registered users is also cumulated within a global webmap thus allowing every registered user to contribute to the progressive setup of collaborative knowledge base about the use of the web.

4. RELATED WORKS

Other projects have already developed different tools providing users with graphical and dynamic environments to play and share their data, bookmarks, history files, etc.

One of the first of such projects is the **Internet Cartographer** from Inventix Software (available for older Windows operating systems and unfortunately not maintained anymore). This software was designed to manage user's navigation history, bookmarks, co-occurring links and had tools like research by keywords or tags. However, Internet Cartographer did not address the issue of information sharing between peers.

Since 2009, Firefox users can install **Voyage**, a plugin developed by Hsiao-Ting Yu. Voyage is designed to map and visualize navigation history, to provide users with an interactive tool that enables them to search into their history files.

Voyage's users can "travel" within a visual web history navigation map organized along a timetable. They can select a specific date, see visit frequencies and co-frequencies and display visited multimedia files on a Media Wall. They can also integrate their twitter feeds. However, generated data cannot be neither shared, nor edited.

Delicious is another social bookmarking website for storing, sharing and discovering web bookmarks. It offers the possibility to its members to share their bookmarks using non-hierarchical classification system of tagged bookmarks with freely chosen index terms (thus generating a kind of folksonomy). Delicious allows its users to compare bookmarks and share their interests, but it does not provide any visual environment to explore the available data.

TouchGraph Navigator 2 is a commercial tool, available since 2001, that provides its users the possibility to visualize and interact with information. With the *TouchGraph Navigator Desktop* users can perform the analysis and create a network-like visualization of their personal data, either on their local computer or on a local server. *TouchGraph Navigator Web* hosts the visualization on a web server.

TouchGraph displays relations between information units, translates raw data into intuitive visualizations, and allows to share information and export images. However, unlike Buzzaar, it is a commercial tool that requires a license and it does not provide support for collaborative knowledge building.

Finally, it is important to mention that the Buzzaar project is also coupled with the **SPADS project** , a collaboration between EPFL and the University of Neuchâtel, aiming at the design of distributed, peer-to-peer techniques allowing privacy preserving transmission of data to aggregating servers. In this perspective, the long term goal of the Buzzaar project is to integrate the techniques developed in SPADS into the Buzzaar toolbar in order to strengthen the aspect of the project related to privacy protection.

5. CONCLUSION

As web users we spend significant time on-line, browsing for different types of information. This activity is already partially stored by our browsers in our “navigation logs” and therefore generates data that could be interesting and useful if made easily understandable and exploitable. The resulting information would then also represent an interesting knowledge to share with peers and collaboratively explore.

Within the Buzzaar project we have designed an experimental prototype that aims at fulfilling these goals. The Buzzaar webmap manager is a freely accessible Java applet allowing users to visualize their web site visit frequency and co-frequency data in the form of an interactive, intuitive and easy to use graphical representation, the webmap. It also provides operations (e.g. filtering, tagging, merge and projection) allowing users to easily manipulate their webmaps in order to better understand and exploit their informational content. The Buzzaar prototype also provides an infrastructure to share webmap data to collaboratively build global webmap representative of larger user communities. We believe that the Buzzaar initiative, if accepted by users, is an interesting step towards the emergence of a new web-based, participative culture, where interactions of individual users with the web can be progressively and collaboratively transformed into a meaningful and exploitable knowledge available for any web user.

6.REFERENCES

Bibliography

Bauwens, M., 2005., "Peer to Peer and Human Evolution: On "the P2P relational dynamic" as the premise of the next civilizational stage" Available from <http://integralvisioning.org/article.php?story=p2ptheory1 > [Accessed 16.2.2011]

C. Wallis, J.C., 2005., "Databases & Data Visualization: the State of the Art, Center for Embedded Networked Sensing", UCLA, Data Management Team

Felber, P., Rajman,M., Riviere,E., Schiavoni, V. and Valerio.J.,2010.,”SPADS: Publisher Anonymization for DHT Storage”in Proceedings of P2P 2010 (10th IEEE International Conference on Peer-to-Peer Computing), Delft, The Netherlands

Klanten, R. Bourquin, S.E.N.,Tissot,T., 2010.,Data Flow: Design graphique et visualisation d'information, Gestalten, Berlin

Lau, A. Moere,A.V., 2007., "Towards a Model of Information Aesthetics in Information Visualization", IEEE International Conference on Information Visualisation (IV'07), IEEE, Zurich , Switzerland, pp. 87-92 Available from <http://web.arch.usyd.edu.au/~andrew/publications/iv07.pdf >[Accessed 16.2.2011]

Manovich, L.,2002.,"The Anti-Sublime Ideal in Data Art", Manovich.net, Available from <http://www.manovich.net/DOCS/data_art.doc > [Accessed 16.2.2011]

Manovich,L.,2010., "What is Visualization?" Manovich.net, Available from <http://manovich.net/2010/10/25/new-article-what-is-visualization/ > [Accessed 16.2.2011]

"Network Collaboration: Peer To Peer As A New Way Of Living - Video Interview with Michel Bauwens",Available from <http://www.masternewmedia.org/news/2006/09/29/network_collaboration_peer_to_peer.htm >

[Accessed 16.2.2011]

Web-references:

"About the LIA", Available from < http://liawww.epfl.ch/ >[Accessed 16.2.2011]

"Aggregation", The Free Dictionnary, Available from < http://www.thefreedictionary.com/aggregation > [Accessed 16.2.2011] "Creative Commons: About",Creative Commons, Available from < http://creativecommons.org/about >[Accessed 16.2.2011]

Delicious, 2003, by Joshua Schachter, Available at < http://www.delicious.com > [Accessed 16.2.2011]

Internet Cartographer, 1999, Inventix Software (version 2.0.1 freeware 2005 Plateform: Windows 95,98, NT), Available at < http://www.inventixsoftware.com/ >[Accessed 16.2.2011]

"Meetopia: collage webzine & digital culture", Metopia.net, Available from <http://www.meetopia.net/index.php?id=about> [Accessed 16.2.2011]

P2P Foundation Contributors, 2011, "The Foundation for P2P Alternatives", p2pfoundation.net, Available from < http://p2pfoundation.net/ > [Accessed 16.2.2011]

TouchGraph Google Browser, 2001, TouchGraph, LLC , Available at < http://www.touchgraph.com/navigator > [Accessed 16.2.2011]

Visual Complexity web site , Available from < http://www.visualcomplexity.com > [Accessed 16.2.2011]

Voyage, 2009, by Hsiao-Ting Yu, Firefox add-ons, Available at <https://addons.mozilla.org/en-US/firefox/addon/voyage/ >[Accessed 16.2.2011]

Webopedia, "What is granularity?",Webopedia , Available from <http://www.webopedia.com/TERM/G/granularity.html > [Accessed 16.2.2011]

Wikipedia contributors, 2011, "Co-occurance", in Wikipedia, The Free Encyclopedia , Available from <http://en.wikipedia.org/wiki/Co-occurrence > [Accessed 16.2.2011]

Wikipedia contributors, 2011, "Data granularity", in Wikipedia, The Free Encyclopedia , Available from <(http://en.wikipedia.org/wiki/Granularity#Computing > [Accessed 16.2.2011]

Wikipedia contributors, 2010, "Distributed knowledge", in Wikipedia, The Free Encyclopedia , , Available from <(http://en.wikipedia.org/wiki/Distributed_knowledge > [Accessed 16.2.2011]

Wikipedia contributors, 2011,"Peer-to-peer", Wikipedia in Wikipedia, The Free Encyclopedia , , Available

from <http://en.wikipedia.org/wiki/Peer-to-peer > [Accessed 16.2.2011]

Wikipedia contributors, 2010, “The commons”, in Wikipedia, The Free Encyclopedia , Available from <http://en.wikipedia.org/wiki/Commons > [Accessed 16.2.2011]