

IBM Research Report

TagOrbitals: A Tag Index Visualization

Bernard Kerr
IBM Research Division
One Rogers Street
Cambridge, MA 02142



Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

TagOrbitals : a tag index visualization

Bernard Kerr
IBM Research
bernard_kerr@us.ibm.com

Abstract

With the rise of web-based tagging systems, it has become important to help users search and browse the index of tags that are being created. One promising approach is visualization, but existing techniques rely on hard-to-interpret network diagrams or focus only on a small subset of a collection. This sketch introduces *TagOrbitals*, a technique that can display an entire index at once, while simultaneously showing details of individual tags and items.

Introduction

Tags are becoming a popular way for people to categorize and organize web-based content such as bookmarks, blog posts, or photos. Tags are simple freeform keywords that are used to add metadata to an object. In most of these tagging systems, an object can be tagged with one or more of these keywords, which results in the object being indexed in multiple categories. For example, a website might be bookmarked with the tags: “blog”, “visualization” and “design”. After many of these objects have been tagged, a *tag index* accumulates. These tags point back to the original objects, and, as more objects are tagged the index becomes larger and more complex. Therefore, new ways of presenting this index are needed to help people search and browse these collections.

Previous visualization techniques have been designed to show keyword or category hierarchies [Shneiderman et al. 2000; Hearst 1999]. But one of the distinctive characteristics of these new tag indexes is that they are not hierarchical. This is because the freeform nature of tag entry does not enforce any structure or fixed taxonomies. More recently, others have tried to visualize these new tag indexes via complex network diagrams, which are hard to interpret (see Figure 1). Or alternatively, they have chosen to selectively present small subsets of tags alone, which do not give users a broad overview of the index (see Figure 2).

Another common way to show a tag index is the “TagCloud”. A TagCloud is an ordering of tags (typically alphabetical) that have been scaled or colored according to the frequency a tag has been used within the collection of objects. Examples of TagClouds can be seen on popular websites such as del.icio.us [<http://del.icio.us>], Technorati [<http://www.technorati.com>] and Flickr [<http://flickr.com>] (see Figure 3). A key quality of a TagCloud is its compactness. To achieve this, it sacrifices other important information such as relationships between tags.

TagOrbitals described below take the opposite approach sacrificing compactness to simultaneously show all the tags and their relationships to each other as well as summary information about the objects that have been tagged. To illustrate this technique, we have constructed a TagOrbitals visualization of one user’s del.icio.us bookmarks.

Method

TagOrbitals are loosely based on the Bohr model of the atom. Each tag is placed in the center (the primary “nucleus” position)

of a series of concentric circles (the “orbits”). Every other tag that has been used in conjunction with the primary tag is placed in a surrounding orbit or band. Each band level indicates the number of other tags used for each bookmark item. The first band shows all the tags that have been used solely with the primary tag. The second band contains any set of two tags that have been used together with the primary tag, and so on (as seen in Figure 4). The title of each bookmark in the collection is drawn radially beside each tag set on each band, so bookmarks that have been classified with the same set of tags have their titles sitting next to each other. Thus, one can easily spot clusters of bookmarks, as seen in Figure 5 “design, graphic”. As in a TagCloud, each TagOrbital (“atom”) is then scaled in proportion to the number of times that its primary tag has been used. The TagOrbitals are laid out using an efficient packing algorithm that places the smallest TagOrbitals in the lower left corner and the largest (most used tags) TagOrbitals in the upper right corner (see Figure 6).

Results

The TagOrbitals visualization can be used for any tag index. An individual user’s index constitutes a portrait of their use of the system, and can be examined and compared to other users. One can scan and zoom into each TagOrbital within the visualization and see the relationship between tags and the actual title of every bookmark within the collection. The number of bands within a TagOrbital shows, at a glance, the number of other tags that have been used with its primary tag. Many bands indicate that the tag is used with many other tags. Few bands means that the tag is used alone. A cluster of bookmark titles on each band shows a focus of interest.

Future Work

The current implementation of TagOrbitals lays the foundation for a new way of visualizing tag indexes. There are several natural directions for future work. The layout of the tags on each band and the individual TagOrbitals themselves could convey additional information about the content and relationships. For example, TagOrbitals with tags in common could be placed close together. The visualization could be made interactive and ultimately dynamically connect to a working system allowing users to navigate to the actual tagged content through the visualization. A final promising direction would be to add temporal data to show how a tag index has evolved over time.

References

- HEARST, M., 1999. The Use of Categories and Clusters in for Organizing Retrieval Results, *Natural Language Information Retrieval*, Strzalkowski (ed.), Kluwer Academic Publishers.
- SHNEIDERMAN, B., FELDMAN, D., ROSE, A., XAVIER F. G. 2000, Visualizing digital library search results with categorical and hierarchical axes. *ACM conference on Digital libraries*, 56-66

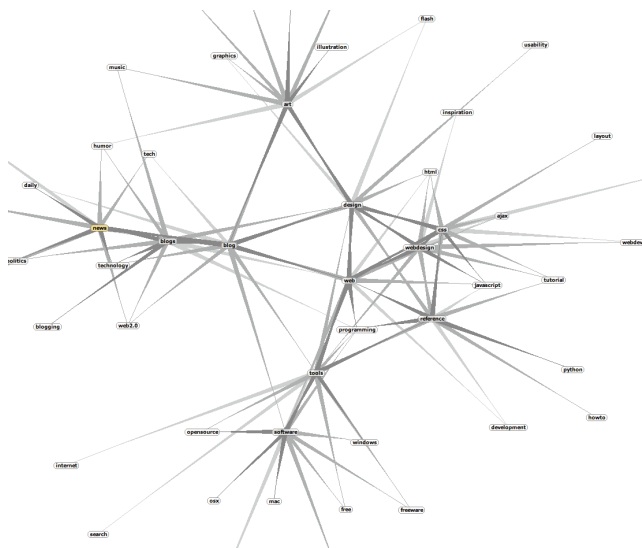


Figure 1. Network graph of delicious tags.
<http://hublog.hubmed.org/archives/001049.html>

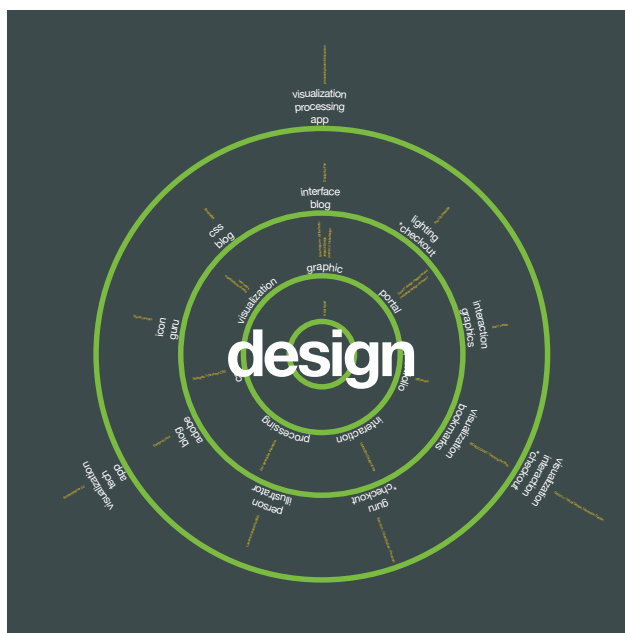


Figure 4. TagOrbital for the “design” tag. The three circular bands, show every tag set associated with “design”. The first band displays single tags, the second band has pairs of tags, and the third band has sets of three tags.

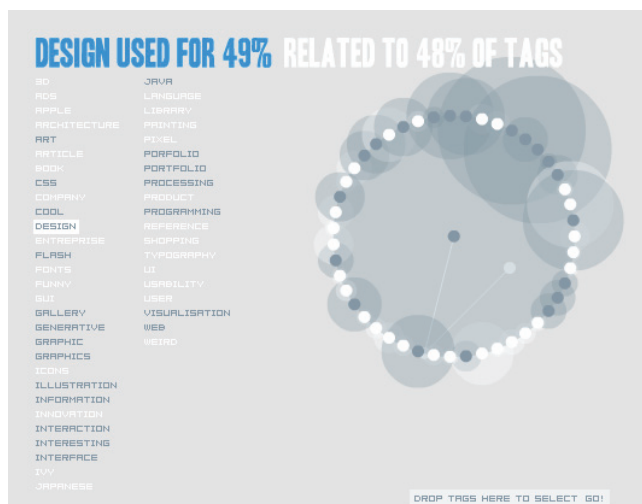


Figure 2. Visualization of a subset of tags from del.icio.us. “Design” tag in the center with related tags on circumference.
<http://www.ivy.fr/revealicious/>

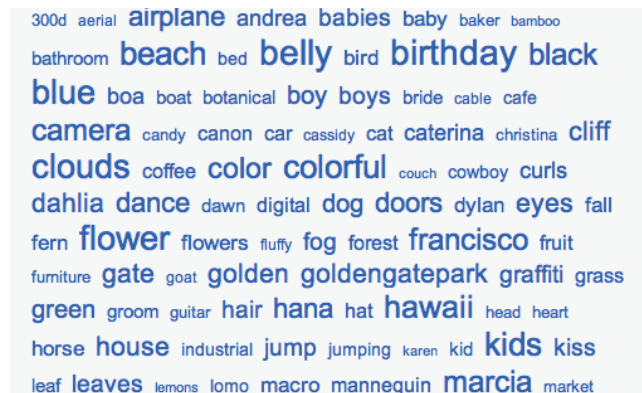


Figure 3. TagCloud that shows tags which are scaled in proportion to how frequently they are used (sorted alphabetically).
<http://www.flickr.com/photos/tags/>



Figure 5. Close up of a TagOrbital for the “design” tag, showing some tags for band one (“graphic” and “portal”) and tags pairs for band two (“interface, blog”, “lighting, *checkout”, and “interaction, graphics”). The small radial lines of text are bookmark titles. Small clusters of these titles (which use the same tag) can be seen for the “graphic” and “portal tags.

