Tightly Coupled Geospatial Interfaces for Collaborative Systems and Just-in-Time Information Retrieval Agents

Arno Scharl

Know-Center and Graz University of Technology
Knowledge Management Institute; Inffeldgasse 21a, 8010 Graz, Austria
scharl@tugraz.at

Abstract. Recent advances in collaborative Web technology are governed by strong network effects and the harnessing of collective intelligence through customer-self service and algorithmic data management. As a result, information spreads rapidly across Web sites, blogs, Wiki applications, and direct communication channels between members of online communities who utilize these services. These various media can be integrated by means of ontology-based tools for building and maintaining contextualized information spaces. The complexity of these spaces calls for new interface technologies that enable users to switch between semantic and geospatial topologies with ease. This paper reviews the literature and outlines the application of geospatial technology for building collaborative systems, and for presenting the results of just-in-time information retrieval agents in an intuitive manner. It introduces *Knowledge Planets* as a new interface metaphor that leverages the new generation of geobrowsing platforms such as NASA World Wind and Google Earth as a front-end for semantic services (see http://www.idiom.at/).

Keywords: Tightly Coupled Interface, Geospatial Projection, Collaborative System, Knowledge Planet, Information Retrieval, Annotation Service

1 Introduction

Integrating cartographic data with geo-tagged knowledge repositories, the emerging Geospatial Web "may ultimately be the big disruptive innovation of the coming decade" [1, xxv]. As such, it will catalyze change and enable a broad range of as yet unforeseen applications. The following hypothetical scenario outlines how the envisioned integration of contextualized information spaces and geospatial technology could radically change individual working environments in the not too distant future, impact the workflow within and across organizations, and enrich the interaction between content providers and their target audience:

Kathryn O'Reilly is a knowledge worker who sells her ability to gather, filter and prioritize electronic content. In a virtual world built on contextualized information spaces, Kathryn seamlessly switches between geographic and semantic topologies. She begins her typical working day floating in the virtual space above earth, ready to *navigate the globe and semantic structures via subtle movements of her eyes, and to access an extensive portfolio of add-on functionality through haptic devices. From her elevated position, Kathryn not only observes the rise and decay of topics, but also the unfolding of social structures based on the unique social networks of her friends and business contacts. Across these networks she builds and shares her knowledge repository, and composes media products that are continuously being validated and enriched by the latest news feeds and third-party multimodal sources.*

The underlying content management system automatically tailors the format of her articles to the preferences of her regular readers. Kathryn adds, selects, categorizes, aggregates, filters and extrapolates information along multiple dimensions, with minimal cognitive requirements. She can structure her daily workflows, access archives of historic textual and multimedia data, and customize her virtual environment with various communication services to interact with pre-defined or dynamically assembled groups of like-minded individuals. At any point in time, Kathryn may use portions of the information space to initialize what-if scenarios and advanced economic or scientific simulations, investigating the complex interplay between computer-generated and real-world participants.

2 Annotation Services

The unprecedented success of the Web 2.0 and geo-browsing platforms has popularized the process of "annotating the Planet" [2]. Both semantic and collaborative technologies unfold their full potential through network effects and benefit from a critical mass of annotations [3]. At present, however, many metadata initiatives still suffer from the chicken and egg problem of wishing that existing content was retrofitted with metadata [4]. This "capture bottleneck" results from the beneficiaries' lack of motivation to devote the necessary resources for reaching a critical mass of metadata [5]. Manual annotation remains difficult, time consuming and expensive [6]. Automated document enrichment [3, 7] addresses this shortcoming by parsing existing Web resources and annotating content fragments along multiple dimensions.

2.1 Geospatial Annotation

Once geospatial context information becomes widely available, any point in space will be linked to a universe of commentary on its environmental, historical and cultural context, to related community events and activities, and to personal stories and preferences. There are several sources of geospatial context information for annotating knowledge repositories:

- Annotation by the authors, manually [8] or through location-aware devices such as GPS navigation systems, RFID-tagged products, and cellular handsets. These devices geo-tag information automatically when it is being created.

- Determining the location of the server – e.g. by querying the Whois[1] database for domain registrations, monitoring how Internet traffic is routed, or by analyzing the URL for additional cues [4].
- Automated annotation of existing documents. The processes of recognizing geographic context and assigning spatial coordinates are commonly referred to as *geo-parsing* and *geo-coding*, respectively [4].

2.1 Extraction and Disambiguation of Geospatial Context

Electronic resources contain metadata as explicit or implicit geographic references. This includes references to physical features of the Earth's surface such as forests, lakes, rivers and mountains, and references to objects of the human-made environment such as cities, countries, roads and buildings [9]. Addresses, postal codes, telephone numbers, and descriptions of landmarks also allow to pinpoint exact locations [4, 10].

At least 20 percent of Web documents contain easily recognizable and unambiguous geographic identifiers [11]. News articles are particularly rich in such identifiers, since they generally report on the location where an event took place, or where it was reported from [12]. The BBC article "Vienna Marking Mozart Milestone" [13], for example, has a target geography of EUROPE/AUSTRIA/VIENNA, and a source geography of EUROPE/UNITED KINGDOM/LONDON. In addition to target and source geography [14], natural language processing can also be used to extract the geographic scope (= intended reach) of Web resources [15].

Named entity recognition locates and interprets phrasal units such as the names of people, organizations, and places [16, 17]. As with most named entity recognition tasks, ambiguity, synonymy and changes in terminology over time complicate the geo-parsing of Web documents [14, 18, 19]. The heterogeneity of spatial references remains a challenge for even the most sophisticated algorithms. Identical lexical forms may refer to distinct places with the same name (VIENNA referring to the capital of Austria as well as a town in Northern Virginia, US), or can have both geographic and non-geographic meanings. TURKEY denotes both a large gallinaceous bird and a bi-continental country between Asia and Europe. Geo-parsing services need to correctly process references to identical or similar places that may be known under different names, may belong to different levels of the administrative or topographical hierarchy, or may be nearby by some measure of proximity [9].

2.1 Assigning Geospatial Coordinates

Once a location has been identified, the documents can be assigned precise spatial coordinates – latitude, longitude, and altitude – by querying structured geographic indices (gazetteers) for matching entries [20, 21]. This process of associating documents with a formal model is also referred to as "document enrichment" [5, 7]. Examples of formal models in the geographic domain are the Geographic Names Infor-

[1] http://www.whois.net/

mation System,[2] the World Gazetteer,[3] the classifications of the United Nations Group of Experts on Geographical Names,[4] the Getty Thesaurus of Geographic Names,[5] and the ISO 3166-1 Country Codes.[6]

While simple gazetteer lookup has the advantage of being language-independent, more advanced algorithms consider lexical and structural linguistic clues, as well as contextual knowledge contained in the documents; e.g. dealing with ambiguity by removing stop-words, identifying references to people and organizations [22], and applying contextual rules like “single sense per document” and “co-occurring place names indicate nearby locations”. For each identified reference, this process assigns a probability *P(name, place)* that a given name refers to a particular place [14]. The location that receives the highest probability is then assigned a canonical taxonomy node such as EUROPE/AUSTRIA/VIENNA; 48°14’ N, 16°20’ E.

2.1 Managing Geospatial Context

Ontologies play a key role in managing geospatial context. While conflicting definitions of “ontology” abound [23], most researchers agree that the term refers to a designed artifact formally representing shared conceptualizations within a specific domain [24, 25]. Deriving ontologies from unstructured textual resources [26] helps validate and extend domain-specific ontologies. Special emphasis should be placed upon integrating unstructured and structured information sources, and developing generic services encompassing both types of information. Ontologies support query term expansion and disambiguation, relevance ranking, and Web resource annotation. Visualizations of ontological structures can also serve as a navigational aid for the more experienced users.

Geo-ontologies encode geographical terms and their semantic relationships – e.g. containment, overlap, and adjacency [20]. In the case of spatially aware search engines, ontological knowledge supports query term expansion and disambiguation, relevance ranking, and Web resource annotation [27]. Geo-ontologies can either be expressed with the same markup language (e.g., OWL DL) as the domain ontology, or use the more specific Geographic Markup Language (GML) of the Open Geospatial Consortium [28].

2 Emerging Interface Technologies

Academia and industry alike call for advanced navigation and interaction facilities to access complex knowledge repositories, following the information seeking mantra of Shneiderman [29] and applying techniques from the emerging field of visual analytics [30]. A new generation of geospatial interface technologies addresses this call, facili-

[2] http://geonames.usgs.gov/
[3] http://www.world-gazetteer.com/
[4] http://unstats.un.org/unsd/geoinfo
[5] http://www.getty.edu/research/conducting_research/vocabularies/tgn/
[6] http://www.iso.org/iso/en/prods-services/iso3166ma/index.html

tating the access and manipulation of geospatial data. Online services such as *Google Maps*, *Yahoo! Local Maps*, and *MSN Virtual Earth (Windows Live Local)* are evolving quickly, adding new functionality, data sources and interface options in rapid succession. In the category of truly three-dimensional representations, dubbed the "holy grail of mapping" [31], the scale-independent spherical globes of Google Earth[7] and NASA World Wind [32][8] dominate the market. These geo-browsers aggregate and project layers of metadata onto cartographic displays.

Effectively navigating complex information spaces requires new interface metaphors. In conjunction with geospatially referenced information spaces, geo-browsers can be used to map annotated content units, track the virtual movements of individual users, and visualize the structure and dynamics of virtual communities. Diverting them from their traditional purpose and coupling them with semantically referenced information spaces, they can also be used to visualize and navigate "Knowledge Planets" based on layered thematic maps. Two distinct interface metaphors, the Earth and knowledge planets, allow users to rapidly switch between geospatial and semantic topologies, providing an unprecedented level of transparency and control over complex and heterogeneous datasets. The ease of switching between interface metaphors and the underlying context acknowledges the increasing importance of finding associations between information entities.

One-dimensional interfaces of text-based search engines and simple navigational systems are insufficient to visualize contextualized information spaces, as they would overwhelm users with an abundance of linear data. Besides, a reduction in dimensionality always goes hand in hand with a loss of information (e.g. projecting a 100-dimensional term space onto a two-dimensional visualization space) – the number of target dimensions positively correlates with the precision of the visualization.

Visualizing social interactions and the rapid diffusion of information increases the challenge of finding a suitable interface metaphor. As the concepts of "desktop", "village" and "landscape" have shown, well-known interface metaphors are powerful instruments to gain market acceptance [33]. Interface metaphors exploit the implicit processing capacity of the human's visual sense. They take advantage of the user's experience from previous perceptions [34], and her ability to navigate in two or three dimensions, recognize patterns, track movements, and compare objects of different size and color [35].

2.1 Geospatial Projections

NASA World Wind and Google Earth promote the "planet" metaphor by providing users with an accurate visual representation and allowing them to browse geospatial data from a satellite perspective. Most geo-browsers offer Application Programming Interfaces (APIs) or XML scripting to encourage developers of third-party online services [36]. Multiple layers of icons, paths and images can be projected via these services. Various visual elements are scaled, positioned on the globe, and linked to the resources of the contextualized information spaces such as Web documents, photo

[7] http://earth.google.com/
[8] http://worldwind.arc.nasa.gov/

collections, and personal contacts. A good example of using NASA World Wind to display additional layers of information is the Moderate Resolution Imaging Spectroradiometer (MODIS) service, providing daily updated planetary imagery at resolutions up to 250-meters-per-pixel, and documenting natural events such as fires, floods, storms, and volcanic activity.[9] The left screenshot of Figure 2 shows a MODIS overlay of Hurricane “Katrina” as of 29 August 2005.

The availability of APIs is largely responsible for the growing popularity of location-based services, often implemented as a mash-up that combines publicly available data and interface services from more than one provider into an integrated user experience [37]. The map in the center of Figure 2 displays the original Sigalert.com service that aggregates real-time traffic data from the San Francisco Bay Area. The screenshot on the right uses Sigalert.com data to visualize accidents and current traffic speeds onto the Google Earth representation of Southern Los Angeles.[10]

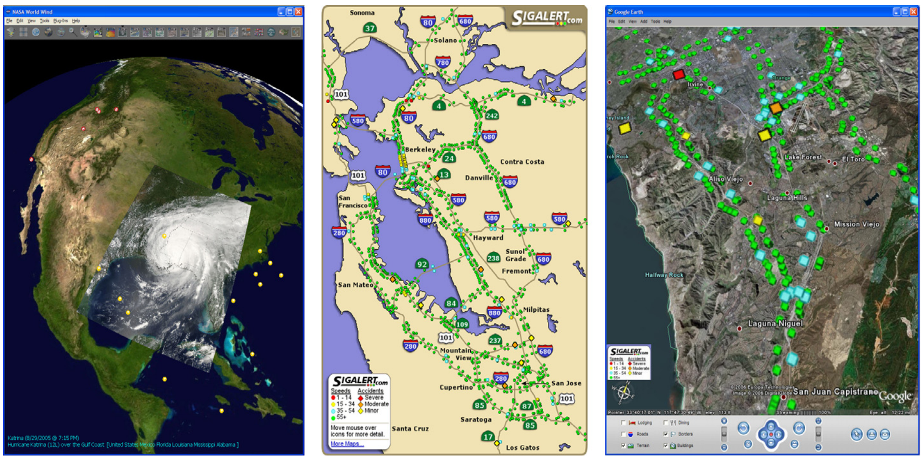

Fig. 1. Data Integration with NASA World Wind and Google Earth

2.2 Geo-temporal News Browsers

Hybrid models of individual and collaborative content production are particularly suited for geo-browsers, which can integrate and map individual sources (monographs, commentaries, blogs), edited sources (encyclopedias, conference proceedings, traditional newsrooms), evolutionary sources (Wiki applications, open-source project documentations), and automated sources (document summarizers, news aggregators). Geo-browsing technology not only impacts the production of content, but also its distribution, packaging, and consumption. When specifying preferences for personalized news services, for example, geo-browsers are effective tools to pinpoint locations and specify geographic areas to be covered.

[9] http://modis.gsfc.nasa.gov/

[10] http://bbs.keyhole.com/ubb/download.php?Number=75329

Personalized news services rely on content fragments annotated along semantic, spatial and temporal dimensions. The geographical scope of an article allows filtering and prioritizing electronic content in line with the user's current task, which can differ from her actual location. Topical similarity is another common dimension to tag and filter news content, often matched against user-specific degree of interest functions. Finally, by adding a temporal dimension (e.g. through time distribution graphs or visual animation), change over time along any other dimension can be captured: the unfolding of events, news distribution patterns, or the inter-individual propagation of messages. Dynamic queries, interactive time displays and playback controls will enable users to identify the rise and decay of topics – e.g. the diffusion of news coverage on natural disasters, or the popularity of specific tourism destinations.

The news summary on the left side of Figure 2 gathers the News Feeds of Associated Press,[11] processes them with the Yahoo! Geocoding API,[12] and displays the results via the Google Maps interface.[13] The second screenshot of Figure 2 shows a geo-temporal news browser that allows users to search a news database via query terms and time-interval sliders, and presents matching articles mapped onto a region of interest [38]. It follows Shneiderman's [29] information seeking mantra: provide an overview, allow zoom and filter, and present details on demand. These guidelines avoid clutter in the display, which results from projecting too many content items from a large knowledge repository simultaneously [19]. Instead of showing the complete set of available news items, for example, a user may wish to restrict the display to articles on climate change that were published in the online editions of Italian newspapers within the last 48 hours.

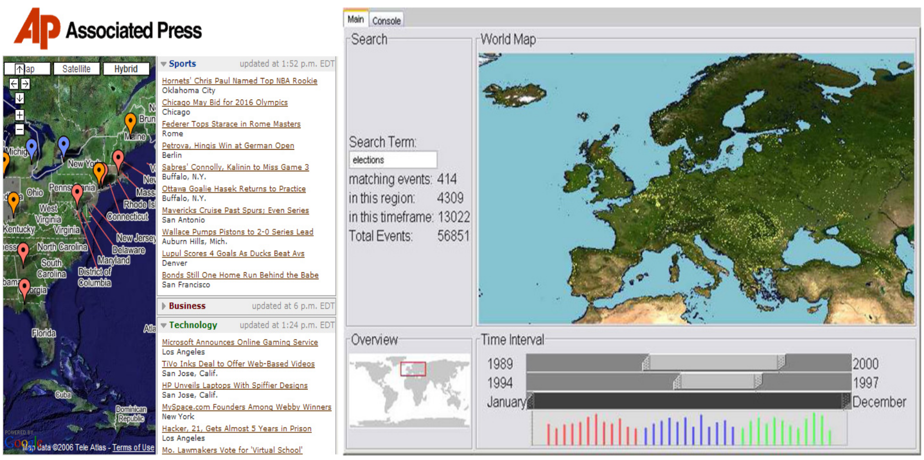

Fig. 2. Interfaces for Accessing Geo-Referenced News Archives

[11] http://hosted.ap.org/dynamic/fronts/RSS
[12] http://developer.yahoo.com/maps/rest/V1/geocode.html
[13] http://81nassau.com/apnews/

2.3 Knowledge Planets

Geo-browsers not only allow visualizing geographic data, but can also serve as a generic image rendering engine to project other types of imagery. Generally, two sets of information need to be integrated and mapped to latitude and longitude: image tiles and terrain information. Using information landscapes instead of cartographic data as the data source for the mapping process represents a radically new way of using geo-browsing technology. Implementations of thematic maps, i.e. visual representations of semantic information spaces based on a landscape metaphor [39], are visually compelling but often did not succeed beyond proof-of-concepts or early prototypes. The restriction to two dimensions and the significant overheads involved in developing and testing truly interactive client applications often kept researchers from disseminating their ideas to a wider audience.

Knowledge planets address both limitations. They can be generated by tiling thematic maps and orthographically projecting them onto the spherical globes of geo-browsers. When defining graphical excellence, Tufte recommends graphical displays that reveal the data at several levels of detail. He identifies layering and separation as the most powerful devices for reducing noise and enriching the content of displays [40]. The layered structure of knowledge planets, analogous to Landsat-7 data of different resolutions, follows this recommendation and allows visualizing massive amounts of textual data. The peaks of the virtual landscape represent focused media coverage on a particular topic, whereas valleys represent sparsely populated parts of the information space. Zooming provides an interactive, intuitive way of selecting the level of aggregation. Unique resource identifiers link concepts embedded in the thematic maps to external sources such as news articles, encyclopedia entries, and papers in scientific journals.

At the time of map generation, the knowledge planet's topology is determined by the structure and content of the knowledge base. Applying the planet metaphor, search results can be visualized as cities, landmarks, or other static objects of the man-made environment. With such a query interface, accessing and filtering complex data along multiple dimensions is as intuitive for users as using a geo-browser to get a glimpse of their next holiday destination. Knowledge planets hide the underlying complexity of the contextualized information space.

Previous research at Graz University of Technology has resulted in VisIslands [41], an algorithm for thematic mapping similar to SPIRE's Themescape [42] and its commercial successor Cartia/Aureka.[14] The algorithm, which supports dynamic thematic clustering of documents, can be adapted to serving image tiles for geo-browsing platforms. The knowledge repository is first pre-clustered using hierarchical agglomerative clustering [43]. The cluster centroids are then distributed randomly in the viewing rectangle. The documents belonging to each cluster, as determined by the initial pre-clustering, are then placed in a circle around each centroid. This arrangement is fine-tuned using a linear iteration force-directed placement algorithm adapted from Chalmers [44]. Shown in Figure 3, the result resembles a contour map of islands. Fortunately, algorithms based on force models easily generalize to knowledge planets' spherical geometries.

[14] http://www.cartia.com/static/aureka.htm

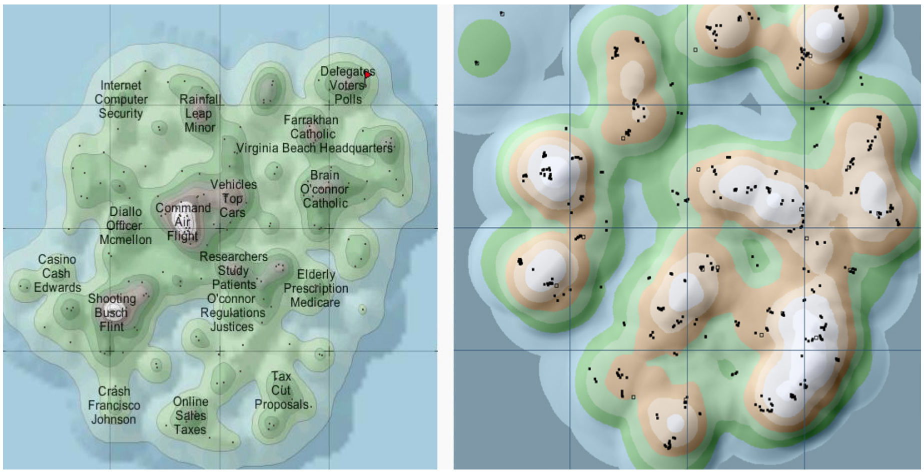

Fig. 3. Information Landscapes of Cartia/Aureka (left) and VisIslands (right)

Thematic mapping algorithms can be used to dynamically cluster and visualize search results while users navigate the contextualized information spaces, with new results being incorporated into the thematic landscape as they arrive. For very large document collections, the topology can be pre-computed and updated on a regular basis, for example overnight [45]. Based on query terms, search results are gathered from the knowledge repository, processed for keywords, clustered, and visualized. When users interactively explore the knowledge planet, they can display and compare additional properties of the search results (document source, date of publication, etc.). To refine their query, they simply have to choose the appropriately labeled clusters in the visualization, providing a quick and convenient way of specifying additional search requests.

The transition from two-dimensional thematic maps to knowledge planets poses a number of conceptual and technical challenges. The spherical globes of knowledge planets complicate the definition of the semantic topology – e.g. the initial arrangement of the major concepts (which will be guided by the domain ontology). Users will expect a consistent experience when rotating the planet. This requires a seamless flow of concepts when crossing the planet's 0° meridian line. The same principle applies to zooming operations. Multiple layers of thematic maps in different resolutions have to be synchronized with each other, and with the adaptive relevance ranking algorithms that avoid clutter by determining which and how many icons are incorporated into the display (clutter usually results from attempts to project too many content items simultaneously). Previous research has developed methods to automatically summarize, split or merging labels and other visual elements that tend to produce clutter [45].

On the server side, the number of layers as well as the possible resolution of the most detailed layer will largely be determined by the availability of computational resources. On the client side, geo-browsers gain from the wide-spread availability of high-end graphics cards for spherically projecting textures and vertices.

2 Tightly Coupled Interfaces

For online marketers, failing to respond to user queries with the most relevant results is a missed marketing opportunity. Traditionally, attempts to tackle this problem have focused on enhancing document classification and ranking algorithms. More recently, the field has started to look at methods to help users improve their queries through disambiguation and query refinement. Domain-specific search engines [46] are therefore becoming increasingly popular because they enable to interpret and refine queries with increased accuracy.

Figure 4 shows the sketch of a browser extension, which aims to reduce the cognitive overhead caused by a lack of semantic, spatial and social context [47]. The *detailed view* on the left side displays the currently active document – either in reading mode (e.g. accessing a Web page), or in writing mode (e.g. collaboratively authoring a document). Context-providing *overviews* in the upper right side visualize the position of this document within the knowledge planet's semantic space, and its geographic positioning.

Below the overviews, just-in-time information retrieval agents generate a summary of related documents, as well as a list of other users associated with the document's content. Users of just-in-time information retrieval agents have been found to be more efficient at retrieving information, and to incorporate more information into their daily workflows that they would with traditional search engines [48]. To aid users in remembering their location, a "you are here" indicator is embedded in the overviews in the upper right corner, as well as an ontology-based navigational aid (the small hierarchical diagram on the right side of the screen).

The active document needs to be tagged automatically and in real time. If a user edits the active document, for example, the position of the "you are here" indicators in the overviews should reflect these changes immediately. Geo-tagging allows the browser extension to project the document onto the correct physical location. Semantic tagging allows the browser extension to (i) rotate the knowledge planet accordingly and mark the document's position on its surface; (ii) update the ontology-based navigational aid. Browsing the Web (i.e. changing the detailed view) thus updates the position of the indicator in the overview. Hence, the various views are said to be tightly coupled [49, 50].

Such *tightly coupled views* are straightforward to implement and understand. User actions in either part of the window cause modifications in all the other related display. The continuous and synchronized display of several views on the contextualized information space will allow rapid and reversible interaction. The advantages and positive effects of tightly coupled interface components on user performance have been documented in several studies [51-53]. Tightly coupled interfaces reduce search time, allow the detection of patterns, and aid users in choosing the next node to visit [54, 3, 55].

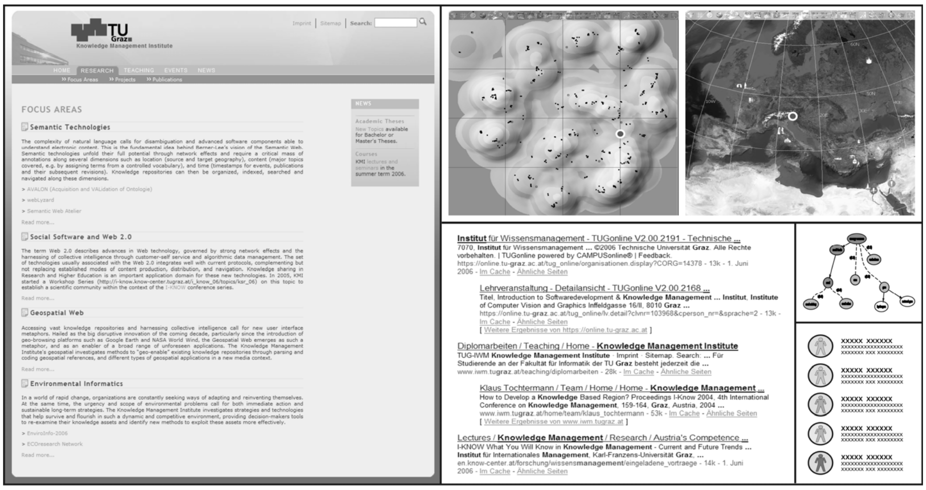

Fig. 4. Sketch of a tightly coupled view showing a rendered document, two distinct overviews (knowledge planet and geospatial projection), a small ontology-based navigational aid, and two views for the just-in-time retrieval of related Web resources and user profiles

The screenshot of Figure 4 is just one possibility of composing various interface services into an integrated user experience on a single 16:9 screen. Modular service architectures allow adapting the layout depending on the requirements of a specific application, or individual preferences. Users not interested in the geospatial context of their activities, for example, can choose to replace the geographic overview with a more detailed rendering of the ontology. For the most demanding users, a multiple screen setup should be most appropriate: the usual desktop environment on the center screen (e.g. document, ontology-based navigational aid, and just-in-time information retrieval agents for related documents, RSS feeds and blogs), and two separate screens for depicting Earth and the knowledge planet, respectively.

3 Visualizing User Activity

Observing, aggregating and visualizing human behavior is a common activity [50], in the past often restricted by the availability of technology. The left diagram of Figure 5 exemplifies the labor-intensive customer tracking typical for traditional retailing outlets in the 1970s (Becker, 1973). Electronic collaborative environments do not need to track the movements of users in a separate process, since their movements and interactions are an integral part of the contextualized information space. These interactions can be regarded an implicit source of user feedback. Explicit feedback is gathered during the interaction in the form of user reviews or recommendations (content, locations, expertise of other users, etc.), and during the formal user interface evaluation, which is independent from the system's actual operation.

As outlined in the preceding section, the knowledge planet’s topology is determined at the time of map generation. Similar to Certau’s *Wandersmänner [56]*, users represent a collective and virtual presence hovering above this topology. Their movements can be traced on the knowledge planet's surface in terms of their paths (here well-trodden, there very faint) and trajectories (going this way and not that). Knowledge planets are excellent platforms for visualizing and analyzing individual and aggregated user interactions. Not only can they show where users are located at the moment, but also where they have already been during the current or previous visits (their "footprints" or "trails"). Extensions like the specific form and direction of footsteps, or dust clouds raised by recent steps should be explored (inspiration can be taken from the study of animal behavior, ants for example, and from various entertainment industry products). Comparing the user's own path with the paths taken by others represents a special type of collaborative filtering and provides reassurance when implemented as a supplemental navigational system.

Analytic services should consider such social information, complementing currently used ranking algorithms that focus on content production, either via document similarity metrics or the number of incoming links. Content consumption can be incorporated as an important indicator of perceived importance, similar to Amazon's collaborative filtering functionality to identify related books, but applied to generic content items. The incorporation of social behavior also resonates well with the Web 2.0 philosophy, and the idea of harnessing collective intelligence.

The same principle not only applies to the individual users of the system, but also to the analysts interested in macro-level effects like the overall patterns in search behavior, or the most popular concepts in a given timeframe. Aggregated representations such as Telegeography's Global Internet Map[15] shown in the middle of Figure 5, however, are often static in nature. User behavior tends to fluctuate heavily, and thus requires more dynamic visual metaphors. Weather maps seem an ideal candidate, as they are a proven method to condense highly complex and dynamic datasets into a visual representation that is useful for experts and understandable for general public at the same time. Techniques developed for weather maps could be used for both local and global trends in the users' interactions. Clouds might represent their aggregated search behavior, for example, while isobars highlight areas of similar interest. Animated isobars would be ideal to visualize trends in content consumption. Those "iso-interest contours" would connect resources sharing the same level of user interest. Clear and sunny high pressure areas would contain topics of limited interest to the average user, while turbulent low pressure systems would indicate heightened levels of activity. Thunderstorms would symbolize the current hotspots of user activity. Many of these hotspots would be triggered by intensive media coverage on exogenous events, and subsequent discussions; e.g. natural disasters such as last year's hurricane "Katrina", and the recent earthquake which struck central Java. Other hotspots might emerge through a process of resonance, when insignificant events or social interactions trigger massive reactions within the contextualized information space.

This opens up interesting analytical opportunities, when potential causes of observed user behavior are mapped against that behavior over time. Isobars could represent real-world data, and color-coding the users' behavior. Then the isobars most closely resembling the color-coding would be most likely to be responsible for the trend in question.

---

[15] http://www.telegeography.com/

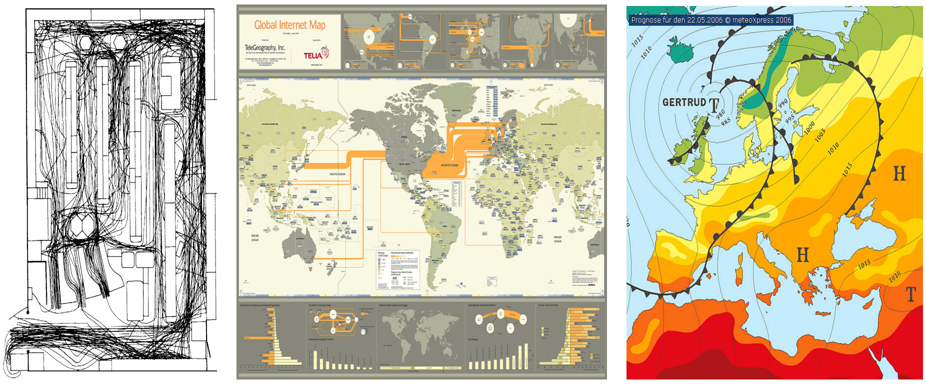

Fig. 5. Traditional customer tracking in a retailing outlet (left); Telegeography[16] Global Internet Map (middle); MeteoXpress[17] weather map as of 22 May 2006 (right)

5 Summary and Outlook

The geo-tagging of existing resources is rapidly gaining the attention of researchers and practitioners alike, because geography emerges as a fundamental principle for structuring complex knowledge repositories, yielding the world's knowledge through the lens of location [31, 58]. Adding location metadata to existing resources and using geo-browsers to access these resources weds physical and virtual spaces. This integration deepens our experiences of information spaces and incorporates them into our every-day lives [36].

This paper presented a conceptualization of advanced collaborative technology based on content, annotation and ontology services to build contextualized information spaces, interface technologies that let users choose between semantic and geospatial topologies for intuitively navigating these spaces, and a framework for analyzing content diffusion and interaction patterns within interactive environments.

References

[1] Erle, S., Gibson, R., Walsh, J.: Mapping Hacks - Tips & Tools for Electronic Cartography. O'Reilly, Sebastopol (2005)

[2] Udell, J.: Annotating the Planet with Google Maps. InfoWorld March 04, 2005 (2005) http://www.infoworld.com/article/05/03/04/10OPstrategic_1.html

[3] Benjamins, R., Contreras, J., Corcho, O., Gómez-Pérez, A.: Six Challenges for the Semantic Web. AIS SIGSEMIS Bulletin 1 (2004) 24-25

[4] McCurley, K. S.: Geospatial Mapping and Navigation of the Web. In: Proc. 10th International World Wide Web Conference (2001) 221-229

[16] http://www.telegeography.com/
[17] http://en.meteoxpress.com/

[5] Motta, E., Shum, S. B., Domingue, J.: Ontology-Driven Document Enrichment: Principles, Tools and Applications. International Journal of Human-Computer Studies 52 (2000) 1071-1109
[6] Ciravegna, F., Dingli, A., Petrelli, D., Wilks, Y.: User-System Cooperation in Document Annotation Based On Information Extraction. In: A. Gomez-Perez and V. R. Benjamins, (eds.): Knowledge Engineering and Knowledge Management - Ontologies and the Semantic Web (Lecture Notes in Artificial Intelligence, Vol 2473). Springer (2002) 65-80
[7] Domingue, J., Motta, E.: PlanetOnto: From News Publishing to Integrated Knowledge Management Support. IEEE Intelligent Systems 15 (2000) 26-32
[8] Daviel, A., Kaegi, F. A.: Geographic Registration of HTML Documents (IETF Internet-Draft, July 2003). Internet Engineering Task Force, Sterling (2003)
[9] Jones, C. B., Alani, H., Tudhope, D.: Geographical Information Retrieval with Ontologies of Place. In: D. R. Montello, (ed.) International Conference on Spatial Information Theory: Foundations of Geographic Information Science (= Lecture Notes in Computer Science, vol. 2205). Springer (2001) 322-335
[10] Ding, J., Gravano, L., Shivakumar, N.: Computing Geographical Scopes of Web Resources. In: Proc. 26th International Conference on Very Large Data Bases (2000) 545-556
[11] Delboni, T. M., Borges, K. A. V., Laender, A. H. F.: Geographic Web Search based on Positioning Expressions. In: Proc. 2nd International Workshop on Geographic Information Retrieval (2005) 61-64
[12] Morimoto, Y., Aono, M., Houle, M. E., McCurley, K. S.: Extracting Spatial Knowledge from the Web. In: Proc. Symposium on Applications and the Internet (2003) 326-333
[13] Bell, B.: Vienna Marking Mozart Milestone. (2006) http://news.bbc.co.uk/2/hi/entertainment/4654880.stm
[14] Amitay, E., Har'El, N., Sivan, R., Soffer, A.: Web-a-Where: Geotagging Web Content. In: Proc. 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (2004) 273-280
[15] Wang, C., Xie, X., Wang, L., Lu, Y., Ma, W.-Y.: Detecting Geographic Locations from Web Resources. In: Proc. 2nd International Workshop on Geographic Information Retrieval (2005) 17-24
[16] Cowie, J., Lehnert, W.: Information Extraction. Communications of the ACM 39 (1996) 80-91
[17] Weiss, S. M., Indurkhya, N., Zhang, T., Damerau, F. J.: Text Mining - Predictive Methods for Analyzing Unstructured Information. Springer, New York (2005)
[18] Kienreich, W., Granitzer, M., Lux, M.: Geospatial Anchoring of Encyclopedia Articles. In: Proc. 10th International Conference on Information Visualisation (2006) In Print
[19] Larson, R. R.: Geographic Information Retrieval and Spatial Browsing. In: L. Smith and M. Gluck, (eds.): GIS and Libraries: Patrons, Maps and Spatial Information. University of Illinois (1996) 81-124
[20] Tochtermann, K., Riekert, W.-F., Wiest, G., Seggelke, J., Mohaupt-Jahr, B.: Using Semantic, Geographical, and Temporal Relationships to Enhance Search and Retrieval in Digital Catalogs. In: Proc. 1st European Conference on Research and Advanced Technology for Digital Libraries (= Lecture Notes in Computer Science, Vol 1324) (1997) 73-86
[21] Hill, L. L., Frew, J., Zheng, Q.: Geographic Names - The Implementation of a Gazetteer in a Georeferenced Digital Library. D-Lib Magazine 5 (1999) www.dlib.org

[22] Clough, P.: Extracting Metadata for Spatially-Aware Information Retrieval on the Internet. In: Proc. 2nd International Workshop on Geographic Information Retrieval (2005) 25-30

[23] Guarino, N.: Understanding, Building and Using Ontologies. International Journal of Human-Computer Studies 46 (1997) 293-310

[24] Jarrar, M., Meersman, R.: Formal Ontology Engineering in the DOGMA Approach. In: R. Meersman and Z. Tari, (eds.): International Conference on Ontologies, Databases and Applications of Semantics (Lecture Notes in Computer Science, Vol 2519). Springer (2002) 1238-1254

[25] Gahleitner, E., Behrendt, W., Palkoska, J., Weippl, E.: Knowledge Sharing and Reuse: On Cooperatively Creating Dynamic Ontologies. In: Proc. 16th ACM Conference on Hypertext and Hypermedia (2005)

[26] Liu, W., Weichselbraun, A., Scharl, A., Chang, E.: Semi-Automatic Ontology Extension Using Spreading Activation. Journal of Universal Knowledge Management 0 (2005) 50-58

[27] Abdelmoty, A. I., Smart, P. D., Jones, C. B., Fu, G., Finch, D.: A Critical Evaluation of Ontology Languages for Geographic Information Retrieval on the Internet. Journal of Visual Languages and Computing 16 (2005) 331-358

[28] Lake, R., Burggraf, D., Trninic, M., Rae, L.: Geography Mark-Up Language: Foundation for the Geo-Web. John Wiley & Sons, New York (2004)

[29] Shneiderman, B.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley, Berkeley (1997)

[30] Wong, P. C., Thomas, J.: Guest Editors' Introduction: Visual Analytics. IEEE Computer Graphics and Applications 24 (2004) 20-21

[31] Levy, S.: Making the Ultimate Map. Newsweek 143 (2004) 56-58

[32] Hogan, P., Kim, R.: NASA Planetary Visualization Tool. In: Proc. American Geophysical Union Fall Meeting (2004)

[33] Fidler, R.: Mediamorphosis: Understanding New Media. Pine Forge Press, Thousand Oaks (1997)

[34] Purves, D., Lotto, R. B., Williams, S. M., Nundy, S., Yang, Z.: Why We See Things the Way We Do: Evidence for a Wholly Empirical Strategy of Vision. Philosophical Transaction of the Royal Society of London 356 (2001) 285-297

[35] Le Grand, B.: Topic Map Visualization. In: J. Park and S. Hunting, (eds.): XML Topic Maps - Creating and Using Topic Maps for the Web. Addison-Wesley (2003) 267-282

[36] Roush, W.: Killer Maps. Technology Review 108 (2005) 54-60

[37] Hof, R. D.: Mix, Match, And Mutate: "Mash-ups" – Homespun Combinations of Mainstream Services – are Altering the Net Business Week 3942 (2005) 72

[38] Rüger, S.: Putting the User in the Loop: Visual Resource Discovery. In: Proc. 3rd International Workshop on Adaptive Multimedia Retrieval (2005)

[39] Chalmers, M.: Using a Landscape Metaphor to Represent a Corpus of Documents. In: A. U. Frank and I. Campari, (eds.): Spatial Information Theory: A Theoretical Basis for GIS (Lecture Notes in Computer Science, Vol 716). Springer (1993) 377-390

[40] Tufte, E. R.: Envisioning Information. Graphics Press, Cheshire (1990)

[41] Andrews, K., Guetl, C., Moser, J., Sabol, V., Lackner, W.: Search Result Visualisation with xFIND. In: Proc. Second International Workshop on User Interfaces to Data Intensive Systems (2001) 50-58

[42] Wise, J. A.: The Ecological Approach to Text Visualization. Journal of the American Society for Science 50 (1999) 814-835

[43] Jain, A., K., Murty, M. N., Flynn, P. J.: Data Clustering: A Review. ACM Computing Surveys 31 (1999) 264-323

[44] Chalmers, M.: A Linear Iteration Time Layout Algorithm for Visualising High-Dimensional Data. In: Proc. 7th Conference on Visualization (1996) 127-132

[45] Andrews, K., Kienreich, W., Sabol, V., Becker, J., Droschl, G., Kappe, F., Granitzer, M., Auer, P., Tochtermann, K.: The InfoSky Visual Explorer: Exploiting Hierarchical Structure and Document Similarities. Information Visualization 1 (2002) 166-181

[46] McCallum, A., Nigam, K., Rennie, J., Seymore, K.: A Machine Learning Approach to Building Domain-Specific Search Engines. In: Proc. 16th International Joint Conference on Artificial Intelligence (1999) 662-667

[47] Turoff, M., Hiltz, S. R., Bieber, M., Fjermestad, J., Rana, A.: Collaborative Discourse Structures in Computer Mediated Group Communications. In: Proc. 32nd Hawaii International Conference on System Sciences (1999)

[48] Rhodes, B. J., Maes, P.: Just-in-Time Information Retrieval Agents. IBM Systems Journal 39 (2000) 685-702

[49] Kumar, H. P., Plaisant, C., Shneiderman, B.: Browsing Hierarchical Data with Multi-level Dynamic Queries and Pruning. International Journal of Human-Computer Studies 46 (1997) 103-124

[50] Scharl, A.: Explanation and Exploration: Visualizing the Topology of Web Information Systems. International Journal of Human-Computer Studies 55 (2001) 239-258

[51] Chimera, R., Shneiderman, B.: An Exploratory Evaluation of Three Interfaces for Browsing Large Hierarchical Tables of Contents. ACM Transactions on Information Systems 12 (1994) 383-406

[52] Norman, K., Weldon, L., Shneiderman, B.: Cognitive Layouts of Windows and Multiple Screens for User Interfaces. International Journal of Man-Machine Studies 25 (1986) 229-248

[53] Shneiderman, B., Shafer, P., Simon, R., Weldon, L.: Display Strategies for Program Browsing: Concepts and an Experiment. IEEE Software 3 (1986) 7-15

[54] Nielsen, J.: Multimedia and Hypertext: The Internet and Beyond. AP Professional, Boston (1995)

[55] Card, S. K., Mackinlay, J. D., Shneiderman, B.: Readings in Information Visualization: Using Vision to Think. Morgan Kaufmann, San Francisco (1999)

[56] De Certeau, M.: The Practice of Everyday Life. University of California Press, Berkeley (1983)

Acknowledgement

IDIOM (Information Diffusion across Interactive Online Media, www.idiom.at) is funded by the Federal Ministry of Transport, Innovation and Technology (BMVIT) and the Austrian Research Promotion Agency (FFG) within the strategic objective FIT-IT (www.fit-it.at). The author wishes to thank his colleagues at the Know-Center and at Vienna University of Economics and Business Administration for their valuable feedback on earlier versions of this manuscript. The Know-Center is funded by the Austrian Competence Center program Kplus under the auspices of the Austrian Ministry of Transport, Innovation and Technology (www.ffg.at), and by the State of Styria.