2013 46th Hawaii International Conference on System Sciences

Media Watch on Climate Change – Visual Analytics for Aggregating and Managing Environmental Knowledge from Online Sources

Arno Scharl, Alexander Hubmann-Haidvogel
MODUL University Vienna,
New Media Technology Department
{arno.scharl, alexander.hubmann}@modul.ac.at

Albert Weichselbraun
University of Applied Sciences Chur, Swiss Institute for Information Research
albert.weichselbraun@htwchur.ch

Heinz-Peter Lang, Marta Sabou
MODUL University Vienna,
New Media Technology Department
{heinz.lang, marta.sabou}@modul.ac.at

Abstract

This paper presents the Media Watch on Climate Change, a public Web portal that captures and aggregates large archives of digital content from multiple stakeholder groups. Each week it assesses the domain-specific relevance of millions of documents and user comments from news media, blogs, Web 2.0 platforms such as Facebook, Twitter and YouTube, the Web sites of companies and NGOs, and a range of other sources. An interactive dashboard with trend charts and complex map projections not only shows how often and where environmental information is published, but also provides a real-time account of concepts that stakeholders associate with climate change. Positive or negative sentiment is computed automatically, which not only sheds light on the impact of education and public outreach campaigns that target environmental literacy, but also help to gain a better understanding of how others perceive climate-related issues.

1. Introduction

The *Media Watch on Climate Change* (MWCC) enhances climate change knowledge and builds awareness about the interdependency of ecological, economic and social issues [39]. By providing integrated access to relevant information sources, MWCC supports the shift from conflict positions to shared meaning, and from isolated knowledge to a shared understanding of system complexity, uncertainty and risk [20; 31]. The integrated information space of MWCC allows investigating the social construction of meaning. Scientists, policy makers, NGOs, corporations, and news media are confronting climate change with an increasing sense of urgency. An enhanced capacity to track the online coverage of these stakeholders helps decision makers to understand public attitudes and behavioral influences as they evaluate possible interventions.

The remainder of this paper is structured as follows: Section 2 outlines the goals and motivation behind MWCC. Section 3 describes the portal and its capabilities to aggregate and manage environmental knowledge from various sources. Section 4 then introduces the Web intelligence platform that MWCC is built upon, followed by a general overview of the user interface and the synchronization mechanism employed to continuously update the portal in Section 5. A detailed description of portal features in Sections 6-8 covers the management and tracking of topics, visual means to investigate the knowledge archive (using media coverage on Rio+20, the *United Nations Conference on Sustainable Development* [41], to exemplify the various methods), as well as query capabilities and data services. Section 9 outlines the chosen systems development approach including evaluation. Section 10 summarizes educational and entrepreneurial opportunities of MWCC technology and its analytical capabilities, before Section 11 concludes the paper and discusses research opportunities in environmental communication, collaboration and decision making.

2. Goals and Motivation

MWCC aims to contribute to a more equal distribution of knowledge, and to support participatory decision making. The following sections outline the goals and motivation of MWCC, structured in terms of application and stakeholder group.

Environmental Education through educational institutions, advocacy organizations and the media are indispensable for achieving sustainability [2]. MWCC provides access to credible information required by environmental educators [15]. Underpinned by science and technology as providers of such information, environmental education is a powerful tool for understanding natural and social processes and their complex interrelationships [21].

1530-1605/12 $26.00 © 2012 IEEE
DOI 10.1109/HICSS.2013.398

Science Communication. Environmental scientists have superior expertise in their focal activities and specific means of disseminating information among their peers. Breaking away from “the mono-logical habits of entrenched and specialized disciplines” [19], however, often remains a challenge. Lack of awareness regarding available scientific expertise remains a problem. MWCC represents a rich document repository for describing the complex phenomenon of climate change [33]. The aim is to increase transparency and give coherence to masses of scientific information [4; 5; 16].

Public Outreach. By providing feedback on outgoing and incoming flows of information, MWCC can improve the effectiveness of public outreach programs, which depend on the quality, professional representation and credibility of communicated content. A longstanding collaboration with NOAA, the *National Oceanic and Atmospheric Administration* [40], emphasizes this aspect of MWCC technology and shows that gathering feedback from social networking platforms can complement established methods of public opinion research. Such social sources offer a unique window into environmental attitudes, lifestyles and behaviors of large user groups. They shed light on individual decision making and document how opinions on environmental issues develop. The structured analysis of social network data helps pinpoint relevant opinion leaders to address when aiming to reach a specific target group.

Stakeholder Coordination. For sharing environmental concerns and responsibilities, individuals and organizations often lack coordination. The competition for budget, jurisdiction, and influence increases their insularity. MWCC addresses this problem by revealing stakeholder interests and enabling businesses, government, and civil society to scrutinize each other in collaborative, consensus-building processes [8]. MWCC aims at improving the quality of decisions, building trust in institutions, and helping resolve conflict among competing interest [10].

Corporate Sustainability. Environmental collaboration platforms can spark interest in redistributing environmental costs more fairly throughout society. Fuelled by increasing environmental awareness and the still unresolved global financial crisis that hit many economies in 2008, there is a growing trend towards accountability, transparency and stakeholder engagement across all levels, functions and operations [13]. This trend stimulates stakeholder interest in the sustainable allocation of corporate resources.

3. Managing Environmental Knowledge

Acquiring, managing and applying knowledge are crucial steps in addressing environmental issues effectively, and ensuring that change is conceived and implemented on both regional and society-wide scales [4]. Climate change is a good example, characterized by diverse opinions of globally distributed stakeholders with different backgrounds and expertise. Understanding the reach of the topics discussed and the opinions voiced by various parties is a complex task that requires knowledge on how specific topics and stakeholders relate to each other.

The MWCC user interface [39] shown in Figure 1 addresses this task. It provides powerful analytical and visual methods to support different types of information seeking behavior such as browsing, search, trend monitoring and visual analytics.

The system detects and tracks emerging topics that are frequently mentioned in a given data sample (typically, an archive of Web documents crawled from relevant sources). The advanced data mining techniques underlying the platform extract a variety of contextual features from the multidimensional document space. A portfolio of synchronized visualizations allows an overall insight into the evolution of the dataset along the dimensions defined by these contextual features (temporal, geographic, semantic, and attitudinal), and subsequent drill-down functionalities to analyze details of the data itself. A key strength of the interface is its use of *multiple coordinated views*, also known as linked or tightly coupled views in the literature [14], where a change in one of the views triggers an immediate update of the others (e.g., when a new document is viewed, the maps pan and zoom to the immediate context of this document).

MWCC harvests data from a range of sources including 150 Anglo-American news media sites, blogs, Web 2.0 platforms including Twitter [43], YouTube [45] and Facebook [35], scientific outlets, and the Web sites of environmental organizations and Fortune 1000 companies [36]. At any given time, only a subset of the vast document space is displayed, depending on the selected source, time interval and affective value (e.g., positive news media articles published in Q1/2012).

The system automatically extracts the dominant issues that are discussed in conjunction with a selected topic (e.g., “climate science” as shown in Figure 1) and displays them through a set of charts that show the frequency and sentiment of a topic, as well as the observable level of disagreement among stakeholders. The search results are also mapped on geographic and semantic maps to show the geographic distribution of the coverage (e.g., places most talked about), as well as its semantic context (e.g., number of documents reporting about a specific issue).

The next section outlines the capabilities of MWCC technology for the real-time processing of Web sources, and for extracting actionable knowledge from these sources.

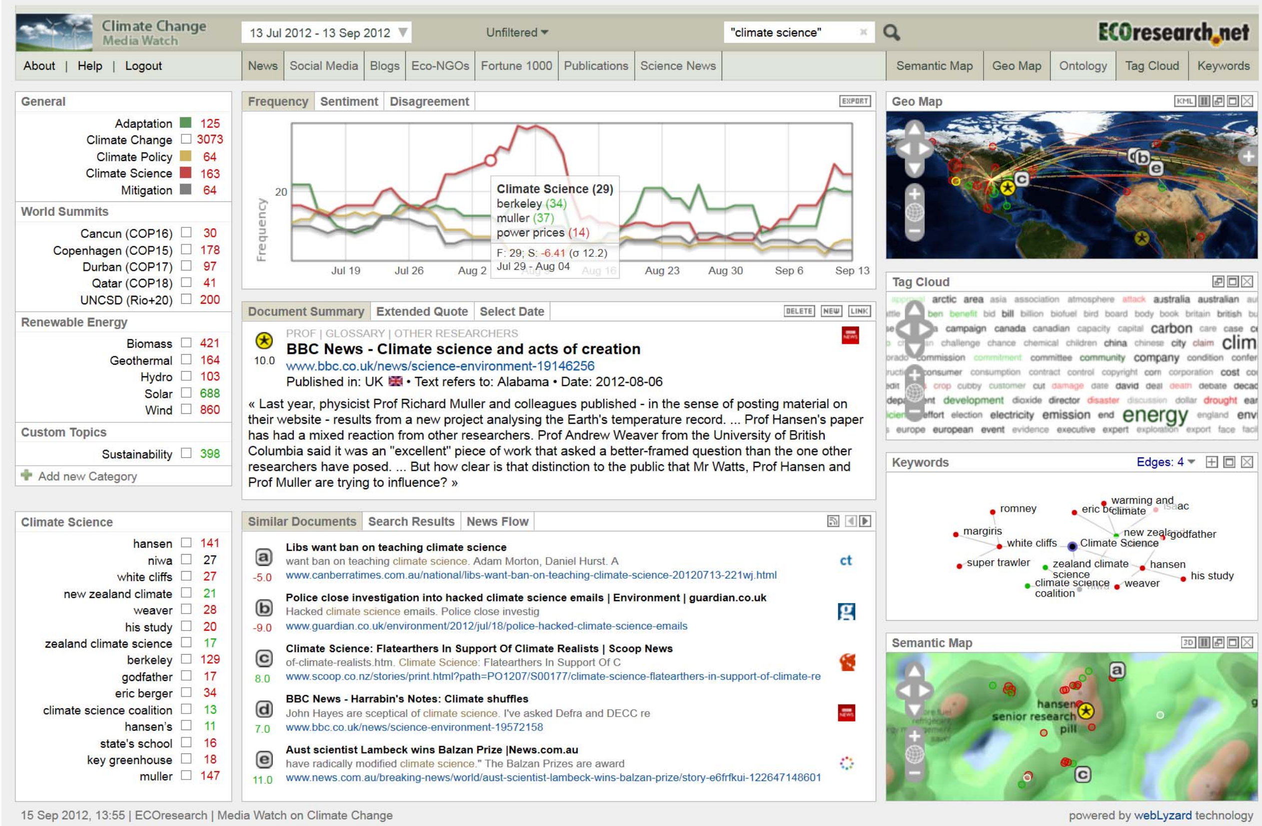

Figure 1. Screenshot of the Media Watch on Climate Change, showing results for a query on "climate science" from Anglo-American news media between July and September 2012 (www.ecoresearch.net/climate)

4. Technology Platform

The goals of MWCC go beyond providing climate-related information – the portal targets stakeholders working on climate change adaptation and mitigation, empowering them to communicate more effectively. Text mining and state-of-the-art semantic technologies capture the human face of big data and allow an unprecedented level of transparency by identifying relevant information from various sources at the touch of a button, including emerging trends and the impact of stakeholder discourse.

To process and enrich streams of information from unstructured, structured and social evidence sources, MWCC utilizes the *webLyzard* media monitoring and Web intelligence platform [44], whose Web crawling and text processing components have proven their scalability and flexibility for many years. webLyzard draws upon expertise from a range of disciplines including human-computer interaction, information visualization, natural language processing and semantic technologies. Such applications shed light on the perceptions of different stakeholder groups, reveal flows of relevant information, and provide indicators for assessing the effectiveness of public outreach activities.

Interactive computing systems have been designed for analyzing social media streams across various domains including sports [18], politics [7; 28; 29] and climate change [14], focusing on specific aspects like (sub-)event detection [1], classification [14] and the analysis of video broadcasts [7]. Such media monitoring tools face two major challenges: (i) collect, analyze and structure very large document collections originating from sources that are heterogeneous in terms of their authorship, formatting, style (e.g., news article versus tweets) and update frequency (weekly, daily or minute-based); (ii) provide an interactive interface to select a relevant subset of the information space, and to analyze and manipulate the extracted data.

The context-sensitive environment presented in this paper allows to analyze and manipulate the extracted knowledge, and to navigate the information space along multiple dimensions. Such an environment, in line with the challenges described above, requires scalable information extraction algorithms and a rapid synchronization of multiple coordinated views. The webLyzard platform provides accurate annotation services to enrich documents with geospatial, semantic and temporal tags. Such annotations describe complex relations, which are best understood in graphical form. For this purpose, the system synchronizes geographic maps, tag clouds, keyword graphs as well as two- and three-dimensional information landscapes. These visualizations help users to understand the context of the extracted knowledge – processing search queries and showing the most relevant documents in their regional context, for example, or comparing the online coverage about an organization by different stakeholders.

5. User Interface Design

The MWCC information exploration and retrieval interface (="dashboard") helps users to interactively identify, track and analyze topics across stakeholders and sources. It builds upon a comprehensive content repository structured along geospatial, semantic and temporal dimensions [26]. The dashboard is divided into six main content areas (see Figure 2):

1. *Sources and Settings:* Drop-down elements in the upper menu let users choose constraints that are relevant for their exploration, including (i) time interval, (ii) document source, and (iii) global sentiment filter (unfiltered, positive, negative). In conjunction with full-text search capabilities and the resulting keyword graph (showing networks of co-occurring terms extracted in real time from the selected source), the global sentiment filter is an effective means to investigate stakeholders associations with a topic.
2. *Topics:* The upper left window of the dashboard serves topic management and content navigation. On mouse-over, users are able to:
   (a) click on a term (representing a topic such as “environment”) to trigger the full-text search;
   (b) use topic markers (small rectangles) to select the topics to be shown in the charts;
   (c) compute related terms to update the associations window ('arrow down' symbol);
   (d) add or modify topics and email alerts with the topic editor ('settings' symbol).
3. *Trend Charts:* Interactive charts show frequency, sentiment and disagreement regarding selected topics in the specified time interval (default: two months).
4. *Content View:* The content view below the trend charts shows the active document including data of publication, keywords, place of publication (source geography), and primary location that is being referenced (target geography).
5. *Search Results:* The lower third of the dashboard displays the list of terms associated with a selected topic, as well as search results on either the document or sentence level).
6. *Visualizations:* To reveal complex and often hidden relations within the document repository, the dashboard integrates geographic maps, ontology graphs, tag clouds, and information landscapes to visualize semantic context. Maps are re-sized automatically and can be repositioned using drag-and-drop operations (see below for a detailed description of the synchronization mechanism).

The visualizations facilitate access to the underlying knowledge base. Clicking on the ‘maximize’ button increases the size of the window; clicking on the ‘popup’ button opens the visualization in a separate browser window (which allows using the system in multiple-screen configurations). Windows can be rearranged by dragging them to the desired position, and (de)activated using the buttons in the top bar.

The various windows are tightly coupled – user actions in one window trigger an immediate update of all other displays – thus supporting an interactive exploration of the information space [14; 25]. As an alternative to entering query terms to find documents, users can click on any position in the maps (not only on the markers) to retrieve articles related to that particular location, topic or domain concept.

Hovering above a map previews the document closest to the current position of the mouse pointer, but does not activate it. When previewing documents, the other visualizations on the right side automatically adjust to show the immediate context of the previewed documents. Users who want to focus on a particular region can disable this default setting and freeze the currently displayed part of the map by pressing the “pause” button.

Users can also adjust the time interval and access historic data via the calendar element. This global setting not only affects the trend chart, as outlined in the next section, but also limits search queries and dynamic visualizations to the chosen time interval.

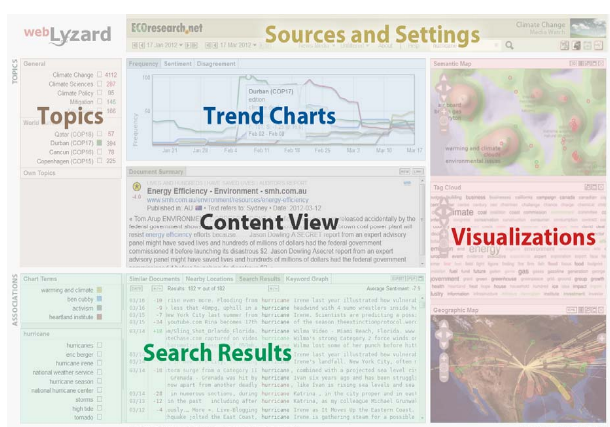

Figure 2. Overview of the main interface elements of the MWCC visual dashboard

6. Managing and Tracking Topics

Registered users can add and modify topics through the topic editor, which also provides the option to set customized email alerts. The 'settings' symbol to activate the topic editor is available next to each topic, implemented as a mouse-over function.

6.1 Topic Editor

Topics are represented as regular expression lists; i.e., term lists with optional wildcards for matching arbitrary character strings. For computing the charts and ranking search results, a document is considered relevant to the topic if it contains at least one of the stated terms. The topic label (= the name to be displayed in the list of topics) itself is not considered in the matching process.

For each topic, users can specify a frequency value that will trigger an alert and choose days to receive the notifications. A sentiment filter allows restricting alerts to positive or negative mentions. Small dots within the topic markers indicate that an active alert exists.

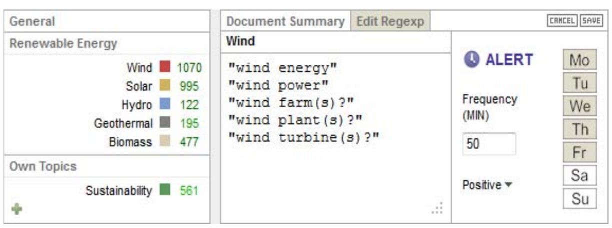

Figure 3. Interface for managing topics (left), editing term lists (middle), and setting e-mail alerts (right)

6.2 Trend Chart

Showing the rise and decay of topics over the last two months (default can be changed via the date selector), the trend chart provides the following time series:

Frequency (total, positive and negative) represents the number of occurrences in the last seven days. Selecting positive or negative coverage through the global sentiment filter affects the data displayed in the chart. Once activated, the footer of the portal also shows the current filter status.

Sentiment shows the average sentiment towards the selected topic for the selected source and interval. The values are based on aggregated polar opinions identified in the document. For achieving accurate results, one needs to deal with the inherent ambiguities of human languages. webLyzard's method to detect sentiment automatically has continually been optimized since 2003 [27], directing particular attention to the context of opinionated terms when resolving such ambiguities [12].

Disagreement computed as standard deviation of the sentiment distribution reflects how contested a particular topic is (the term 'oil spill', for example, tends to have a low standard deviation since everyone agrees on its negative connotation).

Display Features. The vertical axis is rescaled automatically, a feature that is particularly useful if a keyword dominates the coverage and obscures the distributions of other keywords. Hovering above a data point displays the associated keywords and daily statistics (frequency, mean and standard deviation of sentiment). This mouse-over effect identifies topical trends and shows their impact on the individual peaks in the chart. Clicking on a data point triggers a search for this particular topic cluster in the preceding week.

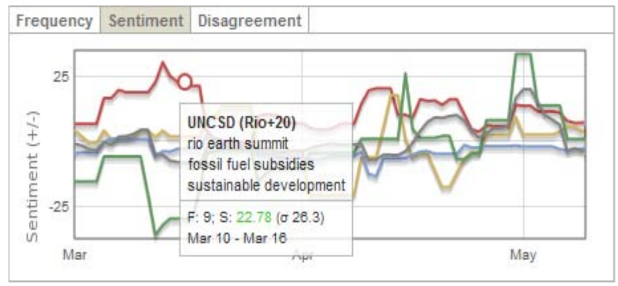

Figure 4. Media sentiment towards major international environmental events (Rio+20, COP15-COP18)

6.3 News Flow Diagram

The interactive news flow visualization shown in Figure 5 reveals emerging topics in the MWCC content archive [6]. It combines several visual metaphors and allows processing multi-source social media datasets as part of an integrated animation. The visualization combines falling blocks with bar graphs and arcs, while keeping these elements clearly separated in different areas of the display. The arc metaphor is adapted and enriched with interactive controls to help users understand the dataset's underlining meaning.

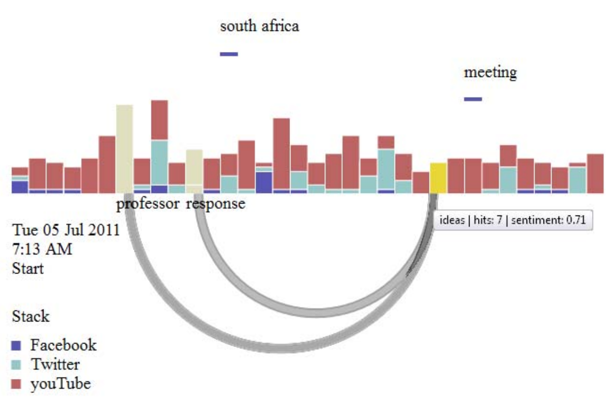

Figure 5. Stacked news flow visualization on Rio+20 social media coverage

7. Visualizing Context and Relations

The right side of the dashboard offers a suite of visualizations that aggregate information along two main context dimensions: semantic and geographic. The *Geographic Map* allows users to interact with the information space in terms of the geographic locations relevant for the documents (i.e., both that of the author and the target of the document). The semantic dimension of the information space is exposed by three different views that leverage increasingly complex semantics: the *tag cloud* is derived from the most frequently mentioned keywords in the information space, the *information landscape* displays clusters of topically related documents thus depicting intrinsic semantic relations between documents, and, finally, the *ontology graph* displays an a-priori constructed semantic model of the domain and assigns each document to the best-matching concept.

7.1 Geographic Map

The geographic map shows the locations of documents based on analyzing their textual content – a process typically referred to as "geo-tagging" [3]. The *Active Document* is highlighted by a yellow asterisk, and the letters [a-e] represent the five highest-ranking documents in the search results view. If interested in a specific location rather than a topic, users can click anywhere on the map to activate the closest document (hovering above the map previews the document, but does not activate it).

After entering a search term, the set of results is visualized in the geographic map. Circular markers show the target geography of the found articles. The diameter of the marker represents the number of matching documents for a given location, its color the average sentiment of these matches.

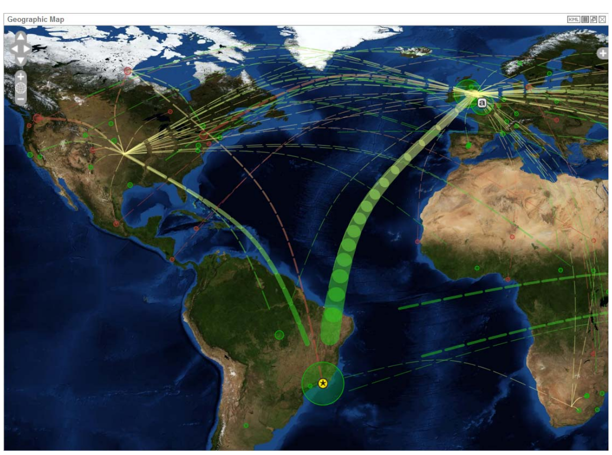

Figure 6. *Geographic Map* revealing global distribution of Rio+20 news media coverage in Q2/2012

Using a color range from yellow to green or red, trajectories link the source and target geography of an article (source = location of publisher; target = main location referenced in the document).

The small '+' symbol on the right side of the window allows selecting alternative base maps (e.g. NASA Blue Marble, Political Borders, Google Terrain, etc.) and deactivating the circular markers or trajectories.

7.2 Tag Cloud

The tag cloud visualizes the most relevant keywords identified in recent online publications of a stakeholder group. Terms are arranged alphabetically. Size and opacity are proportional to their importance; i.e., their relative frequency in the text archive (more frequent terms are rendered in darker shades, using a larger font, less frequent terms using a smaller font and lower opacity settings). Users can adapt the zoom factor of the display and select terms by clicking on them, which triggers a full text search (see Section 8) and displays a list of matching concordances.

The color of terms indicates their sentiment (positive = green; neutral = black; negative = red). This allows investigating the "spin" across sources, as exemplified in Figure 7 – the typically balanced news media coverage (left) compared to more emotional and often negative postings on social media platforms (middle) and the extremely positive slant characteristic for corporate publications found on the Web sites of Fortune 1000 companies (right).

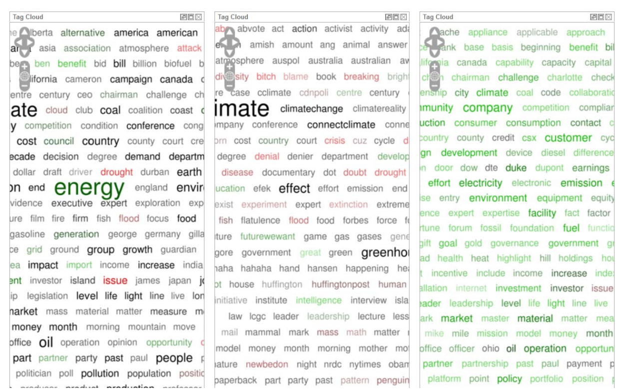

Figure 7. *Tag Cloud* comparison across sources (news media, social media, Fortune 1000), with color-encoded sentiment data

7.3 Information Landscape

To show topical relatedness in large document repositories, information landscapes cluster and visualize massive amounts of textual data [17]. They implement the concept of "location" in an innovative way that transcends the traditional geographic interpretation.

The information landscape resembles a geographic map. Instead of geographic proximity, however, it represents semantic similarity between documents. At the time of map generation, its topography is determined by the content of the knowledge base. The peaks of the virtual landscape shown in Figure 8 indicate abundant coverage on a particular topic, whereas valleys (lighter shades of green) or the ocean (blue) represent sparsely populated parts of the information space.

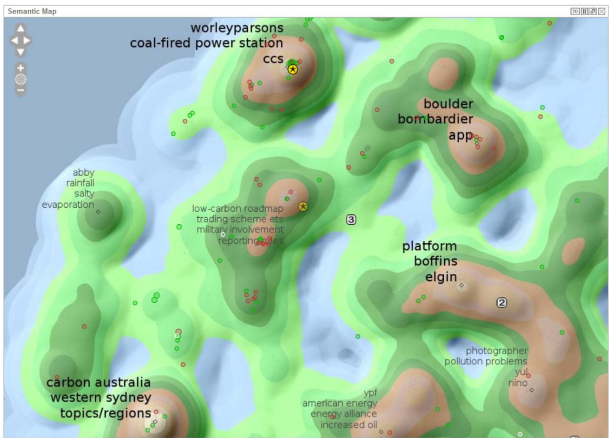

Figure 8. *Information Landscape* with document markers based on a query for "carbon" in Q2/2012

The visualization projects the following interface elements onto the landscape topography [22; 23]:

- *Captions.* The keywords for each peak are calculated based on the content of nearby documents.
- *Document Symbols.* The symbols [a-e] indicate the position of documents from the search results.
- *Search Results* are mapped onto the topography using the same representation as the geographic map (size of the markers indicates the number of results, the color their average sentiment).
- *Document Selection.* Hovering above a location in the map previews the closest matching document in the active document window, clicking selects this document (small gray dots indicate the locations of the individual documents).

7.4 Ontology Graph

Ontologies express shared meaning within a given domain and allow the effective integration of external structured data [32], for example Web-based linked data [30; 38]. The ontology graph of Figure 9 displays a clickable domain model that matches documents and concepts to help users determine their current location in the information space. The domain model used to structure MWCC content has been developed in close collaboration with experts from the Climate Program Office of NOAA, the U.S. *National Oceanic and Atmospheric Administration [40]*.

The ontology graph depicts hierarchical relations as arrows. The active document is highlighted by a yellow asterisk, and the letters [a-e] represent the classification of the five highest-ranking documents in the search results view. Clicking on a concept activates the highest-ranking document for this particular concept.

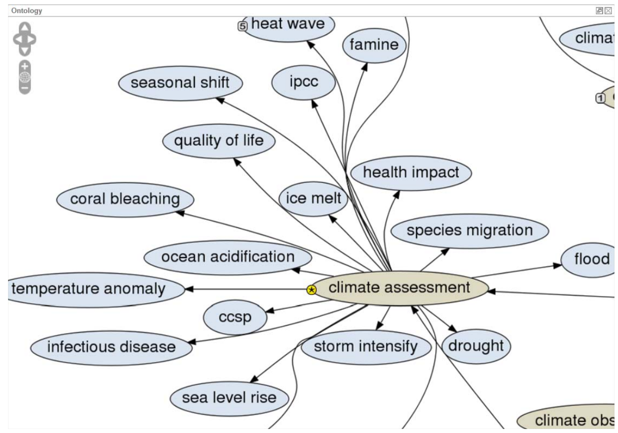

Figure 9. *Ontology Graph* showing the relations between the concept "climate assessment" (grey) and its sub-nodes (blue)

8. Semantic Search and Data Services

The search box is located in the top bar of the portal. The system supports wildcard characters – while the asterisk (*) symbol represents any number of unknown characters, for example, the question mark (?) represents exactly one character. The syntax of additional options (source, location, date, etc.) is described in Section 8.2: "Advanced Search".

A full-text query not only returns the search results, but also updates other views including associations, similar topics, keyword graph, maps as well as the content view, which summarizes the highest-ranking document including its metadata.

8.1 Simple Search

Sentence Level. Upon entering a search query, the system lists all sentences containing the search term and groups them by document. The column headers show publication date, document sentiment, currently selected and total number of matches, and average sentiment of the displayed search results. Users can sort results by their date of publication as well as the sentiment on both the document and sentence level (the color of the term reflects sentence-level sentiment).

Document Level. Once a user selects the document tab, the search results are being shown in a more aggregated representation including document title, URL and favicon of the source. The document symbols (a-e, 1-5) are also used in the various maps to indicate the position of this document. The value underneath the symbols represents document sentiment. Clicking on the text block extends the quote, clicking on the circular symbol activates a particular document. Back and forward buttons allow browsing the list of related documents; the RSS link provides a continuously updated news feed.

8.2 Advanced Search

Additional query options extend the capabilities of the simple search. They enable users to filter the search results by specifying restrictions based on document metadata (title, source, date, geographic location, etc.):

- The documents' title field can be queried with the title: operator; e.g. "carbon title:ocean" will only return documents containing "ocean" in their title.
- The site operator limits the search to a given web domain; searching for "carbon site:cnn.com" returns documents on "carbon" mirrored from the CNN Web site [34].
- Documents in a time interval can be searched using from: and to: operators; searching for "carbon from:2012-06-01 to:2012-06-30", for example, will only return documents published in June 2012.
- The source and target geography of documents can be considered with the source: and target: operators; searching for "carbon dioxide source:US target:Rio de Janeiro", for example, will return documents about carbon dioxide published in the US and referencing the venue of Rio+20 as the main location in the text.

9. Implementation and Evaluation

Following an evolutionary systems development approach [24], rapid feedback cycles and agile software development [9] have been instrumental in MWCC conceptualization and implementation. For this purpose, ongoing usability inspections have been conducted to analyze and assess the system. This low-overhead, heuristic approach asks a team of experts to investigate the interface design against recognized usability principles. The evaluation has been performed periodically during the design and implementation phases, incorporating feedback from NOAA [40] and other partners. Improvements were integrated into the prototype early in the development cycle.

A review and evaluation framework [11; 42] has been developed and applied to the presented platform, comparing its features and metrics with those of other third-party tools. The authors plan to conduct a summative usability evaluation utilizing the eye tracking facilities of the *University of Applied Sciences Chur* [37]. Formal experiments with test users will collect quantitative performance measurements (e.g., time required to successfully complete a given task) and statistically analyze the collected data.

10. Opportunities for Environmental Education and Decision Making

MWCC technology as outlined in the preceding sections creates both entrepreneurial and educational opportunities in the transition to a knowledge-based economy. By uncovering patterns and trends in online media coverage and making them available through an interactive dashboard, the webLyzard platform [44] and its portfolio of semantic technologies helps to optimize communication strategies and reach target audiences effectively. The system addresses the inherent ambiguities of natural languages and guides analysts in specifying queries and configuring the various filter options and output services.

For effective decision making, organizations require such advanced systems for the creation, storage and retrieval of knowledge contained in their heterogeneous systems. The rapid evolution of this knowledge represents an economic imperative for organizations to adopt data integration and "content hub" facilities that provide flexible and real-time access to various sources of digital content. This ensures a continuous supply of relevant information for decision makers, and a common knowledge base for communication among stakeholders within and across organizations. MWCC represents such a "content hub" and provides a unique empirical base to unveil the conditions that lead to the introduction, transfer, and uptake of climate research. Capturing the evolution of climate change knowledge addresses calls to exploit the potential of semantic systems and services in environmental informatics.

By offering a public showcase of the webLyzard social analytics platform, MWCC hopes to pave the way for organizations to adopt semantic technologies by lowering the costs of capturing, integrating and analyzing information. Since user-generated content significantly influences environmental decision making, potential target audiences for webLyzard technology include market research providers, public outreach departments, campaign organizers and policy advisors. These players were among the first to recognize that the Internet is not just a medium to obtain information; it also facilitates communication between individuals who want to share their opinions, perceived threats and shortcomings, and suggestions for change. An increasing number of public platforms enable them to do so, but usually in an unstructured format. Treating such platforms as social evidence sources and transforming their repositories of user feedback into more structured representations provides real value for analysts and public relations experts who increasingly depend on the just-in-time processing of information hidden in public archives.

MWCC potentially serves a broad spectrum of different target groups, from corporations of varying size to non-profit organizations, research centers, government organizations, public policy advisors, and news media outlets. webLyzard's modular portfolio of analytical methods can be embedded into a dedicated Web platform such as MWCC, or be utilized through add-on services for existing applications. The extraction and interactive exploration of knowledge bears significant commercial potential not only for market research and business intelligence applications, but also for campaign management, product development, and monitoring the effectiveness of outreach programs. The gathered information creates feedback loops and shows how well an organization's communication is received, understood, and remembered.

11. Summary and Conclusion

This paper described the *Media Watch on Climate Change* [39], an interactive Web portal to assist environmental decision making by collecting and consolidating climate change knowledge from unstructured, structured and social sources. These sources include the publications of scientists, policy makers, educators, NGOs, news media and corporations – stakeholders who recognize the need for collective action, but differ in their specific worldviews, goals and agendas [4].

These publications are typically being created through processes of cooperation and social exchange. They depend on and benefit from a synergy of skills, distributed decision making, and the dynamic maintenance of shared knowledge. MWCC analyzes this abundance of information by providing automated classification and contextualization services as well as a visual dashboard to access the resulting information space. This increases transparency, assists in the transition from conflict positions to shared meaning, and replaces static and compartmentalized knowledge by the ability to deal effectively with uncertainty, risk and dynamic change.

A well-informed public is more likely to support sustainable policies, and more capable of holding decision makers accountable. MWCC studies public discourse and critical debate that lead to a shared understanding of climate change issues on all political levels, ranging from inter-individual communication and local communities to global campaigns and treaties. By investigating communicative processes between disciplines and stakeholders, MWCC aims to unearth hidden assumptions and misconceptions about climate change, contribute to a mutual understanding of existing problems, and suggest priorities for research and policy development.

12. References

[1] Adams, B., Phung, D., et al. (2011). Eventscapes: Visualizing Events Over Times with Emotive Facets. *19th ACM International Conference on Multimedia (MM-2011)*. Scottsdale, USA: 1477-1480.

[2] Ahlberg, M. (2005). "Integrating Education for Sustainable Development", *Handbook of Sustainability Research*. Ed. W.L. Filho. Frankfurt: Peter Lang. 477-504.

[3] Amitay, E., Har'El, N., et al. (2004). "Web-a-Where: Geotagging Web Content", *27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Sheffield, UK: ACM Press. 273-280.

[4] Bowman, T. (2008). Summary Report: A Meeting to Assess Public Attitudes about Climate Change. Silver Springs: National Oceanic and Atmospheric Administration (NOAA), George Mason University Center for Climate Change Communications.

[5] Bowman, T. (2009). "A Turning Point in Climate Change Communication Priorities", *International Journal of Sustainability Communication,* 4: 64-77.

[6] Braşoveanu, A.M.P., Hubmann-Haidvogel, A., et al. (2012). Interactive Visualization of Emerging Topics in Multiple Social Media Streams. *ACM Working Conference on Advanced Visual Interfaces (AVI-2012)*. G. Tortora et al. Capri, Italy: ACM: 530-533.

[7] Diakopoulos, N., Naaman, M., et al. (2010). Diamonds in the Rough: Social Media Visual Analytics for Journalistic Inquiry. *IEEE Symposium on Visual Analytics Science and Technology*. Salt Lake City, USA: IEEE: 115-122

[8] Doering, D.S., Cassara, A., et al. (2002). *Tomorrow's Markets - Global Trends and Their Implications for Business*. Washington: World Resources Institute.

[9] Dönmez, D. and Grote, G. (2011). "Managing Uncertainty in Software Development Projects", *Agile Processes in Software Engineering and Extreme Programming (= LNCS, Vol. 77)*. Berlin: Springer. 326-328.

[10] Foti, J., Silva, L.d., et al. (2008). *Voice and Choice: Opening the Door to Environmental Democracy*. Washington, DC: World Resources Institute.

[11] Fritzsche, H. (2011). *Social Media Monitoring - A Review and Evaluation Framework for State-of-the-Art Tools and Associated Metrics*. Master Thesis, Department of New Media Technology: MODUL University Vienna.

[12] Gindl, S., Weichselbraun, A., et al. (2010). Cross-Domain Contextualisation of Sentiment Lexicons. *19th European Conference on Artificial Intelligence (ECAI-2010)*. H. Coelho et al. Lisbon, Portugal: IOS Press: 771-776.

[13] Held, S., Lauvergeon, A., et al. (2007). Policy Directions to 2050 - A Business Contribution to the Dialogues on Cooperative Action. Geneva: World Business Council for Sustainable Development.

[14] Hubmann-Haidvogel, A., Scharl, A., et al. (2009). “Multiple Coordinated Views for Searching and Navigating Web Content Repositories”, *Information Sciences,* 179(12): 1813-1821.

[15] Kassas, M. (2002). “Environmental Education: Biodiversity”, *The Environmentalist,* 22(4): 345-351.

[16] Keohane, R.O. and Nye, J.S. (1998). “Power and Interdependence in the Information Age”, *Foreign Affairs,* 77(5): 81-94.

[17] Krishnan, M., Bohn, S., et al. (2007). "Scalable Visual Analytics of Massive Textual Datasets", *21st IEEE International Parallel & Distributed Processing Symposium*. Long Beach, USA: IEEE Computer Society.

[18] Marcus, A., Bernstein, M.S., et al. (2011). Twitinfo: Aggregating and Visualizing Microblogs for Event Exploration. *2011 Annual Conference on Human Factors in Computing Systems*. Vancouver, Canada: ACM: 227-236.

[19] Meppem, T. and Bourke, S. (1999). “Different Ways of Knowing: A Communicative Turn Toward Sustainability”, *Ecological Economics,* 30(3): 389-404.

[20] Oepen, M. (2000). "Environmental Communication in a Context", *Communicating the Environment*. Eds. M. Oepen and W. Hamacher. Frankfurt: Peter Lang. 41-61.

[21] Rose, O.H. and Bridgewater, P. (2003). “New Approaches Needed to Environmental Education and Public Awareness”, *Prospects,* 33(3): 263-272.

[22] Sabol, V. and Scharl, A. (2008). "Visualizing Temporal-Semantic Relations in Dynamic Information Landscapes", *11th International Conference on Geographic Information Science (AGILE-2008), Semantic Web Meets Geospatial Applications Workshop*. Girona, Spain: AGILE Council.

[23] Syed, K.A., Kröll, M., et al. (2012). “Incremental and Scalable Computation of Dynamic Topography Information Landscapes”, *Journal of Multimedia Processing and Technologies*. Forthcoming (Accepted 22 Aug 2012)..

[24] Scharl, A. (2000). *Evolutionary Web Development*. London: Springer. http://webdev.wu.ac.at/.

[25] Scharl, A. (2006). “Tightly Coupled Geospatial Interfaces for Collaborative Systems and Just-in-Time Information Retrieval Agents”, *Research in Computing Science,* 25: 3-18.

[26] Scharl, A. (2007). "Media Platforms for Managing Geotagged Knowledge Repositories", *The Geospatial Web - How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society*. Eds. A. Scharl and K. Tochtermann. London: Springer. 3-14.

[27] Scharl, A., Pollach, I., et al. (2003). "Determining the Semantic Orientation of Web-based Corpora", *Intelligent Data Engineering & Automated Learning (LNCS Vol. 2690).* Eds. J. Liu et al. Berlin: Springer. 840-849.

[28] Shamma, D.A., Kennedy, L., et al. (2009). Tweet the Debates: Understanding Community Annotation of Uncollected Sources. *1st ACM SIGMM Workshop on Social Media (WSM-2009)*. Beijing, China: 3-10.

[29] Shamma, D.A., Kennedy, L., et al. (2010). Tweetgeist: Can the Twitter Timeline Reveal the Structure of Broadcast Events? *ACM Conference on Computer Supported Cooperative Work (CSCW-2010)*. Savannah, USA.

[30] Waitelonis, J., Ludwig, N., et al. (2011). “WhoKnows? - Evaluating Linked Data Heuristics with a Quiz that Cleans Up DBpedia”, *International Journal of Interactive Technology and Smart Education,* 8(4): 236-248.

[31] Wallace, M.G., Cortner, H.J., et al. (1996). “Moving Toward Ecosystem Management: Examining a Change in Philosophy for Resource Management”, *Journal of Political Ecology,* 3: 18-32.

[32] Weichselbraun, A., Wohlgenannt, G., et al. (2010). “Refining Non-Taxonomic Relation Labels with External Structured Data to Support Ontology Learning”, *Data & Knowledge Engineering,* 69(8): 763-778.

[33] Wilson, K.M. (2000). "Communicating Climate Change through the Media - Predictions, Politics and Perceptions of Risk", *Environmental Risks and the Media*. Eds. S. Allan et al. London, New York: Routledge. 201-217.

Online Resources

[34] *CNN International.* http://www.cnn.com/.

[35] Facebook. http://www.facebook.com/.

[36] Fortune 500+. https://www.fortune500-app.com/.

[37] HTW Chur. http://www.cheval-lab.ch/.

[38] Linking Open Data (LOD). http://linkeddata.org/.

[39] MWCC. http://www.ecoresearch.net/climate/.

[40] NOAA. http://www.noaa.gov/.

[41] Rio+20. http://www.uncsd2012.org/.

[42] SOMEMO. http://www.somemo.at/.

[43] Twitter. http://www.twitter.com/.

[44] *webLyzard.* http://www.webLyzard.com/.

[45] YouTube. http://www.youtube.com/.

Acknowledgement

Key components of the system presented in this paper were developed within the DIVINE (www.weblyzard.com/divine) and Triple-C (www.ecoresearch.net/triple-c) research projects, funded by *FIT-IT Semantic Systems* of the Austrian Research Promotion Agency and the *Austrian Climate Research Program* of the Austrian Climate and Energy Fund (www.ffg.at; klimafonds.gv.at).

The authors would like to thank the team of researchers involved in conceptualizing and developing these components: *information landscape* (S. Kamran A. Ahmad, V. Sabol and M. Kröll); *news flow diagram* (A. Brasoveanu); *tag cloud* (W. Rafelsberger, M. Foels); *sentiment analysis* (S. Gindl); *ontology engineering* (G. Wohlgenannt, S. Belk); and *topic management* (T. Wenter).