Scharl, A., Hubmann-Haidvogel, A., Sabou, M., Weichselbraun, A. and Lang, H.-P. (2013). "From Web Intelligence to Knowledge Co-Creation – A Platform to Analyze and Support Stakeholder Communication", *IEEE Internet Computing,* 17(5): 21-29.

From Web Intelligence to Knowledge Co-Creation – A Platform to Analyze and Support Stakeholder Communication

Arno Scharl,[1] Alexander Hubmann-Haidvogel,[1] Marta Sabou,[1] Albert Weichselbraun,[2] Heinz-Peter Lang[3]
[1] MODUL University Vienna, Department of New Media Technology
[2] University of Applied Sciences Chur, Swiss Institute for Information Research
[3] Vienna University of Economics and Business, Research Institute for Computational Methods

Abstract

Organizations require tools to assess their online reputation as well as the impact of their marketing and public outreach activities. The Media Watch on Climate Change is a Web intelligence and online collaboration platform that addresses this requirement. It aggregates large archives of digital content from multiple stakeholder groups and enables the co-creation and visualization of evolving knowledge archives. This paper introduces the base platform and a context-aware document editor as an add-on that supports concurrent authoring by multiple users. While documents are being edited, semantic methods analyze them on the fly to recommend related content. Positive or negative sentiment is computed automatically to gain a better understanding of third-party perceptions. The editor is part of an interactive dashboard that uses trend charts and map projections to show how often and where relevant information is published, and to provide a real-time account of concepts that stakeholders associate with a topic.

1. Introduction

Novel approaches to online collaboration transcend the traditional constraints of time and space, supporting synchronous or asynchronous collaboration in heterogeneous groups. Collaborative tools can be distinguished by the process that they support and their typical user groups. *Collaboratories* [5], for example, are distributed environments that support the curation and production of scientific data by distributed teams of scientists and domain experts. *Crowdsourcing Systems* [4], by contrast, target large numbers of non-expert contributors and invite them to collaboratively address a range of tasks, e.g. editing *Wikipedia.org* articles or annotating images through games with a purpose [1].

Many organizations that require tools to create, manage and retrieve information adopt Wiki-based systems. *Semantic Wikis* allow authors to add metadata to the document, for example to declaratively describe and categorize products and link them to knowledge worker profiles [12]. Such metadata facilitates navigation and supports searching the content base.

This paper presents a Web intelligence and collaborative editing platform that goes beyond Semantic Wikis by automating both the annotation and the retrieval processes. A tight integration between the editor and the analytical components of the Media Watch on Climate Change (MWCC) supports collaborative information seeking. While a document is being edited by a single or multiple users, the content of the authored document acts as an advanced search term over the entire document collection and lets users discover related information. This addresses calls for proactive methods to manage knowledge, helping organizations to capitalize on their employees' skills and expertise in an unobtrusive manner [12]. Using the terminology of Golovchinski [7], the editor enables *implicit* information seeking behavior by "recommending" relevant content along spatial and semantic dimensions, supporting both synchronous and asynchronous information seeking.

The novelty of MWCC lies in blending the features of Semantic Wikis with those of generic collaborative information seeking systems. The recommendations represent an implicit collaborative feature that uses a combination of natural language processing, semantic technologies and visual analytics [3] to retrieve data from a knowledge repository. To build this repository, MWCC collects and analyzes a large number of documents from online sources that are heterogeneous in terms of authorship, formatting and update frequency. It provides a dashboard to select a relevant subset of the information space, and to analyze and manipulate extracted data in real time. This dashboard sheds light on stakeholder perceptions, reveals flow of relevant information between these stakeholders, and provides metrics for assessing the effectiveness of awareness and public outreach campaigns.

To process and enrich the information, MWCC utilizes the *webLyzard.com* media monitoring and Web intelligence platform, whose data acquisition and knowledge extraction services have been optimized for Web-scale applications. These services enrich documents with annotations along multiple dimensions and use these annotations to render interactive visualizations. The visual interface enables users 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, not only requires scalable information extraction algorithms, but also a rapid synchronization of multiple coordinated views. MWCC synchronizes geographic maps, tag clouds, keyword graphs as well as two- and three-dimensional information landscapes. These visualizations help users to understand the specific context of the extracted knowledge.

MWCC technology covers a wide range of application scenarios. It processes search queries to show the most relevant documents in a regional context, uncovers trends in stakeholder communication, measures the success of marketing and public outreach activities, identifies opinion leaders, helps organizations to engage with target audiences, and provides a real-time account of positive and negative topics that these audiences associate with the organizations' products and services.

Building upon previous work describing the Web intelligence aspects of MWCC technology [9], this paper focuses on its knowledge co-creation capabilities of an integrated document editing environment. The remainder of this paper is structured as follows: Section 2 outlines the Web intelligence capabilities of MWCC including the management of evolving domain knowledge from various sources, metrics to track emerging trends, and a visual dashboard to access the knowledge repository. Building upon this Web intelligence infrastructure, Section 3 presents the MWCC document editing environment as a novel approach to create knowledge in a real-time, collaborative manner. Section 4 covers visual means to support the editing process through recommendations based on topical similarity and geographic location. Section 5 describes the iterative and user-centered approach to system development and evaluation. Section 6 summarizes and concludes the paper.

2. Web Intelligence

By capturing stakeholder communication automatically and in real time, Web intelligence technologies allow an unprecedented level of transparency. They identify relevant information from various sources at the touch of a button, and reveal patterns and trends. The domain of *climate change* is well suited to demonstrate the potential of such technologies, as it is characterized by diverse opinions of globally distributed stakeholders with different backgrounds and expertise. These stakeholders need to manage and apply relevant knowledge to address societal issues effectively, and ensure that change is conceived and implemented on both regional and society-wide scales. Understanding the reach of the topics discussed and the opinions voiced by various parties is a complex task that requires knowledge on how topics and stakeholders relate to each other.

Sources. MWCC harvests data from a range of sources including 150 Anglo-American news media sites, blogs, Web 2.0 platforms including *Twitter*, *Facebook*, *Google+* and *YouTube*, scientific outlets, and the Web sites of environmental organizations and Fortune 1000 companies. 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 the first quarter of 2013).

Knowledge Extraction. Automated annotation mechanisms extract various contextual features from the collected documents, including their time stamp and authoring entity, geographic locations being discussed, positive or negative sentiment, as well as a set of keywords to label the document. Automated annotation helps avoid acceptance problems of collaboration tools that require users to categorize knowledge resources manually. It also supports the identification of emerging trends and dominant issues being discussed in conjunction with a selected topic.

Dashboard. The MWCC user interface provides access to the knowledge repository. Search results are mapped onto 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 dashboard's analytical and visual methods support different types of information seeking behavior through six main content elements:

- *Sources and Settings.* The top menu lets users choose constraints that are relevant for their exploration, including time interval to access longitudinal data, document source, and global sentiment filter (unfiltered, positive, negative). These settings not only affect the trend charts, but also limit search results and dynamic visualizations.
- *Topics.* The upper left part of the dashboard provides topic management and content navigation. Users can (i) click on a topic to trigger a full-text search; (ii) use the topic markers (small rectangles) to select topics to be shown in the charts; (iii) compute related terms via the 'arrow down' symbol; and (iv) add or modify topics and email alerts via the 'settings' symbol.

- *Trend Charts.* Interactive charts show weekly frequency, average sentiment, and the level of disagreement regarding selected topics. The sentiment values are based on aggregated polar opinions identified in the document [14]. Disagreement, computed as standard deviation of sentiment, reflects how contested a particular topic is (the term 'oil spill', for example, has a low standard deviation since most people agrees on its negative connotation). Hovering above a data point displays the associated keywords and daily statistics, while a click triggers a search for this topic in the preceding week.
- *Content View.* The content view below the trend charts shows the active document including date of publication, keywords, place of publication (source geography), and primary location that is being referenced (target geography).
- *Search Results.* The full-text search of MWCC supports wildcard characters, Boolean operators, and regular expressions. The results are displayed in the lower third of the dashboard: a list of associated terms, as well as a list of search results with tabs to switch between the document, sentence and source level. Each new query also updates the other windows of the portal.
- *Visualizations:* To reveal complex and often hidden relations within the document repository, the dashboard integrates a portfolio of visualizations to provide insight into the evolution of the underlying document space. The portfolio will be described in Section 4 and includes geographic projections, ontology and keyword graphs, a color-coded tag cloud, as well as an information landscape.

A key strength of the dashboard is its use of *multiple coordinated views*, also known as linked or tightly coupled views in the literature [8], where a change in one of the views triggers an immediate update of the others – while a new document is viewed or edited, for example, the maps pan and zoom to represent its semantic context and offer a holistic, real-time view of the domain. As an alternative to entering query terms for finding documents, users can use the visualizations 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. When previewing documents, the other visualizations automatically adjust to show the immediate context of the previewed documents – a crucial feature to support the knowledge co-creation process outlined in Section 3.

Example: US Election 2012. Integrating the gathered intelligence in campaign management and public outreach applications creates instant feedback loops that show public outreach and communication managers how well their messages are received, understood, and remembered by specific target audiences. By uncovering patterns and trends in online media, the system reveals hidden knowledge and supports decision making. The search query for "Barack Obama" in Figure 1, for example, lists *Mitt Romney* and the *Keystone XL Pipeline* as the two top news media associations in the context of climate change between January and November 2012.

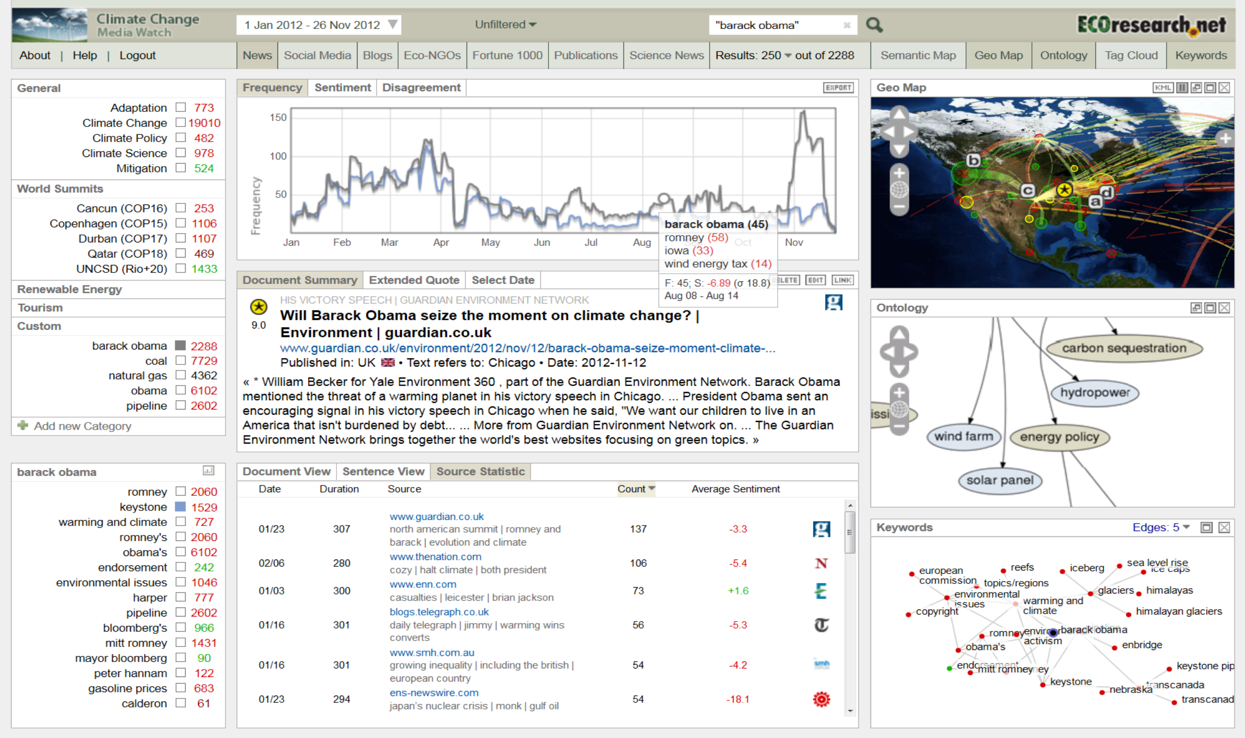
The screenshot also reveals that the topic of climate change has received little attention from campaign managers. This is shown by the strong correlation between the two trend lines. Since coverage on the Keystone XL Pipeline was driven by third parties and not the campaign managers, intentional topics would show up as divergence points (e.g. the minor peak triggered by the *Rio+20 Conference* in June 2012). The notable silence on climate change came to a sudden end once hurricane "Sandy" hit the U.S. East Coast, and New York City Mayor *Michael Bloomberg* endorsed the incumbent for re-election, citing the threat of climate change as the primary reason for his recommendation.

This section has shown how MWCC helps users to manage rapidly changing information shared among stakeholders via a range of channels. Going beyond the analysis of third-party information, the following section describes the potential of knowledge extraction and visualization algorithms for dynamic and collaborative content creation.

3. Knowledge Co-Creation

A new culture of participation is driven by advances in social computing. Many authors consider documents as "open, evolvable seeds rather than finished products" [6]. These documents are created through processes of cooperation and social exchange – depending on and benefiting from a synergy of skills, distributed decision making, and the dynamic maintenance of shared knowledge.

The MWCC platform reflects this trend, supporting explicit and implicit collaborative processes. Explicit support includes (e.1) synchronous editing of documents, (e.2), shared topic definitions via a regular expression editor, and (e.3) exporting jointly created resources for reuse in third-party to support existing workflows of environmental stakeholders. Implicit support is provided through (i.1) customizable content recommendations on the group level through analyzing the co-authored document in real time, and (i.2) the visual methods of Section 4 as adaptive navigational aids that reflect shared meaning and the semantic context of the edited document.



General

Adaptation	☐	773
Climate Change	☐	19010
Climate Policy	☐	482
Climate Science	☐	978
Mitigation	☐	524

World Summits

Cancun (COP16)	☐	253
Copenhagen (COP15)	☐	1106
Durban (COP17)	☐	1107
Qatar (COP18)	☐	469
UNCSD (Rio+20)	☐	1433

Renewable Energy

Tourism

Custom

barack obama	■	2288
coal	☐	7729
natural gas	☐	4362
obama	☐	6102
pipeline	☐	2602

Add new Category

barack obama

romney	☐	2060
keystone	■	1529
warming and climate	☐	727
romney's	☐	2060
obama's	☐	6102
endorsement	☐	242
environmental issues	☐	1046
harper	☐	777
pipeline	☐	2602
bloomberg's	☐	966
mitt romney	☐	1431
mayor bloomberg	☐	90
peter hannam	☐	122
gasoline prices	☐	683
calderon	☐	61

HIS VICTORY SPEECH | GUARDIAN ENVIRONMENT NETWORK

9.0

Will Barack Obama seize the moment on climate change? | Environment | guardian.co.uk

www.guardian.co.uk/environment/2012/nov/12/barack-obama-seize-moment-climate-...

Published in: UK • Text refers to: Chicago • Date: 2012-11-12

« * William Becker for Yale Environment 360 , part of the Guardian Environment Network. Barack Obama mentioned the threat of a warming planet in his victory speech in Chicago. ... President Obama sent an encouraging signal in his victory speech in Chicago when he said, "We want our children to live in an America that isn't burdened by debt... ... More from Guardian Environment Network on. ... The Guardian Environment Network brings together the world's best websites focusing on green topics. »

Date	Duration	Source	Count	Average Sentiment
01/23	307	www.guardian.co.uk north american summit | romney and barack | evolution and climate	137	-3.3
02/06	280	www.thenation.com cozy | halt climate | both president	106	-5.4
01/03	300	www.enn.com casualties | leicester | brian jackson	73	+1.6
01/16	301	blogs.telegraph.co.uk daily telegraph | jimmy | warming wins converts	56	-5.3
01/16	301	www.smh.com.au growing inequality | including the british | european country	54	-4.2
01/23	294	ens-newswire.com japan's nuclear crisis | monk | gulf oil	54	-18.1

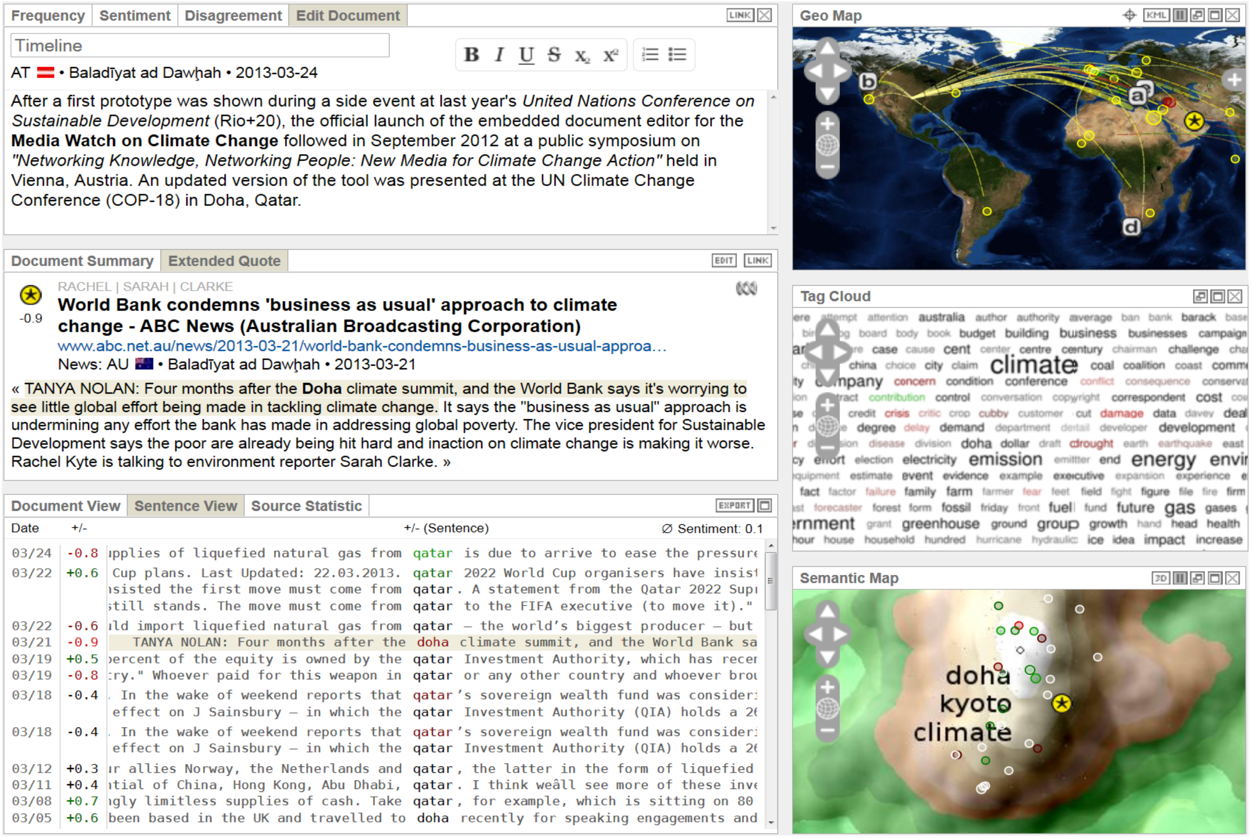
Figure 1. Screenshot of the Media Watch on Climate Change, showing results for a query on “Barack Obama” from Anglo-American news media between January and November 2012 (www.ecoresearch.net/climate)

Applying traditional authoring models to the collaborative creation of documents assumes a sequential process where authors: (i) investigate a topic by reviewing the literature, scanning the latest news media coverage or querying search engines; (ii) communicate with their co-authors; (iii) draft and revise the manuscript in either an asynchronous or synchronous manner [7]. The sequential character of these phases and fragmented workflow is not commensurate with the requirements of today's dynamic markets and agile, globally distributed enterprises.

MWCC’s automated content recommendations are an important step to merging these distinct phases. They go beyond online word processing tools such as *Google Docs* and the *MS Word Web App* by enabling implicit information seeking, where the system infers informational needs from the users’ actions, rather than explicit queries [7]. When multiple users jointly edit a document, the system immediately distributes changes to all co-authors and performs background queries to fetch similar documents from a customizable set of sources (news, social media, NGO Web sites, etc.), and suggests appropriate tags to classify the document (geographic location, sentiment, keywords, etc.). The content recommendations are made on the fly, while users are typing. They help align the evolving document with the existing body of knowledge. Co-authors learn about related threads on social media platforms or recent scientific discoveries. This instant feedback loop can guide the collaborative work to embrace issues that otherwise might have been overlooked.

Given the lack of hyperlinks in many of the edited documents (as compared to those gathered from online sources), MWCC follows a content-based ranking strategy. Slider elements of the advanced search feature allow users to customize these recommendations in line with their preferences and current tasks – i.e., determine the impact of the similarity, geographic proximity and topicality of documents on the ranking process.

Figure 2 shows the collaborative editor in action. Accessible via a separate tab, it is seamlessly integrated with the analytical MWCC components. The screenshot depicts an example of the collaborative editing of a press release about presentations at Rio+20 and COP-18 side events (see Section 5). In the background, the system analyzes the text to suggest related information from the knowledge repository. The top-ranked result is a recent *ABC News* article about the lack of action following COP-18, which the press release’s co-authors can access to refine their arguments accordingly.


Frequency
Sentiment
Disagreement
Edit Document
LINK
B I U S x₂ x²
Geo Map
KML
b
a
d
Tag Cloud
ere attempt attention australia author authority average ban bank barack base
board body book budget building business businesses campaign
case cause cent center centre century chairman challenge
china choice city claim climate coal coalition coast
company concern condition conference conflict consequence
contribution control conversation copyright correspondent cost
credit crisis critic crop cubby customer cut damage data davey deal
degree delay demand department detail developer development
disease division doha dollar draft drought earth earthquake east
effort election electricity emission emitter end energy
estimate event evidence example executive expansion experience
fact factor failure family farm farmer fear feet field fight figure file fire firm
forecaster forest form fossil friday front fuel fund future gas gases
grant greenhouse ground group growth hand head health
hour house household hundred hurricane hydraulic ice idea impact increase
Semantic Map
3D
doha
kyoto
climate

Timeline

AT • Baladīyat ad Dawḩah • 2013-03-24

After a first prototype was shown during a side event at last year's *United Nations Conference on Sustainable Development* (Rio+20), the official launch of the embedded document editor for the **Media Watch on Climate Change** followed in September 2012 at a public symposium on *"Networking Knowledge, Networking People: New Media for Climate Change Action"* held in Vienna, Austria. An updated version of the tool was presented at the UN Climate Change Conference (COP-18) in Doha, Qatar.

Document Summary | Extended Quote | EDIT | LINK

RACHEL | SARAH | CLARKE

-0.9

World Bank condemns 'business as usual' approach to climate change - ABC News (Australian Broadcasting Corporation)

www.abc.net.au/news/2013-03-21/world-bank-condemns-business-as-usual-approa…

News: AU • Baladīyat ad Dawḩah • 2013-03-21

« TANYA NOLAN: Four months after the **Doha** climate summit, and the World Bank says it's worrying to see little global effort being made in tackling climate change. It says the "business as usual" approach is undermining any effort the bank has made in addressing global poverty. The vice president for Sustainable Development says the poor are already being hit hard and inaction on climate change is making it worse. Rachel Kyte is talking to environment reporter Sarah Clarke. »

Document View | Sentence View | Source Statistic | EXPORT

Date	+/-	+/- (Sentence) ∅ Sentiment: 0.1
03/24	-0.8	ıpplies of liquefied natural gas from qatar is due to arrive to ease the pressure
03/22	+0.6	Cup plans. Last Updated: 22.03.2013. qatar 2022 World Cup organisers have insist
		ısisted the first move must come from qatar. A statement from the Qatar 2022 Supı
		still stands. The move must come from qatar to the FIFA executive (to move it)."
03/22	-0.6	ıld import liquefied natural gas from qatar – the world's biggest producer – but
03/21	-0.9	TANYA NOLAN: Four months after the doha climate summit, and the World Bank sa
03/19	+0.5	percent of the equity is owned by the qatar Investment Authority, which has recer
03/19	-0.8	:ry." Whoever paid for this weapon in qatar or any other country and whoever brou
03/18	-0.4	. In the wake of weekend reports that qatar's sovereign wealth fund was consideri
		effect on J Sainsbury – in which the qatar Investment Authority (QIA) holds a 26
03/18	-0.4	. In the wake of weekend reports that qatar's sovereign wealth fund was consideri
		effect on J Sainsbury – in which the qatar Investment Authority (QIA) holds a 26
03/12	+0.3	ır allies Norway, the Netherlands and qatar, the latter in the form of liquefied
03/11	+0.4	ıtial of China, Hong Kong, Abu Dhabi, qatar. I think weâll see more of these inve
03/08	+0.7	ıgly limitless supplies of cash. Take qatar, for example, which is sitting on 80
03/05	+0.6	been based in the UK and travelled to doha recently for speaking engagements and

Figure 2. Collaborative editor showing a document in progress and related content recommendations

The visualizations on the right side of the interface place the current document into its geographic and semantic context. The geographic map shows a stream of related documents referring to Doha, mostly originating from North America. In terms of document content, the press release is matched to a cluster on "kyoto, doha, climate". As the edited document is extended with more content, its similarity to other document clusters is recomputed so the authors can track, in real time, how their document shifts its position in the information space – giving them the opportunity to adjust the content until the document is placed in the vicinity of a desired topic cluster.

Documents created with the MWCC editor become part of the document repository and are being processed together with stakeholder communication collected from external sources – for example, from news and social media. The editor builds upon the *Wikidocs.com* semantic editing framework to support multiple concurrent users, and incorporates the results of a recent research project, the *Climate Change Collaboratory* (www.ecoresearch.net/triple-c). In order to embed the archived knowledge into the existing workflows of environmental stakeholders, MWCC supports a range of export formats including RSS, HTML and PDF for textual data, and CSV for time series date – in line with calls for a *Semantic Social Web*, where data is not locked away within data silos, but can be easily integrated and exchanged between applications and authors [2].

4. Visualizing Related Knowledge Resources

The MWCC dashboard offers a suite of visualizations that aggregate and render information along two main context dimensions, *geographic location* and *topical similarity*. The visualizations provide real-time feedback and support the editing process, as outlined in the following sections.

4.1 Geographic Location

The *Geographic Map* supports three distinct modes to interact with the information space: edit, explore, and search. When rendering the display, the system distinguishes between source and target information – i.e., the locations of the author(s) versus the primary location referenced in a document, which is determined by analyzing its textual content in a process typically referred to as *geo-tagging* [11].

The currently viewed or edited document is highlighted by a yellow asterisk. This represents a particularly useful feature in conjunction with the collaborative editor described in Section 3, as the map automatically pans and zooms to the referenced locations. If interested in a specific location rather than a topic, users can click anywhere on the map to switch from editing to information exploration mode and retrieve the closest document in terms of geographic proximity (hovering above the map previews documents without activating them).

In *search mode* (i.e., after entering a query term), the result set is visualized with circular markers that show the target 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. Trajectories link the geographic source and target of an article.

4.2 Topical Similarity

To show topical similarity in large document repositories, dynamic topography information landscapes cluster and visualize massive amounts of textual data [13]. 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. A peak such as the one shown in the lower right corner of Figure 2 (“doha, kyoto, climate”) indicates abundant coverage on a topic, whereas valleys (lighter shades of green) or oceans (blue) represent sparsely populated parts of the information space. The peak labels are calculated based on the content of nearby documents.

Similar to the adaptive display of the geographic map, the shown part of the topography reflects the content of the currently edited document. Small gray dots indicate the locations of *related documents*. Circular markers, color-coded by sentiment, show the distribution of *search results*.

4.3 Additional Visualization Types

The portfolio of MWCC visualizations contains additional methods to track emerging topics in stakeholder discussions. The portfolio includes *tag clouds*, animated *news flow diagrams,* and graph-based representations of *keywords* and *ontologies* [9]. Two of them are particularly suited to support the MWCC editing environment:

- The **tag** cloud (Figure 2) is a well-known representation and useful navigational aid that is easy to parse for non-expert users. It arranges keywords alphabetically, whose size and opacity are proportional to their relative frequency. The tag clouds of MWCC use color to indicate sentiment (positive = green; neutral = black; negative = red), which allows investigating the "spin" across sources; e.g., balanced coverage of news media compared to emotional and often negative social media postings, or the positive slant characteristic for publications found on corporate Web sites.
- The **ontology graph** (Figure 1) expresses shared meaning within a domain and displays a clickable domain model that matches documents and concepts to help users determine their current location in the information space. Hierarchical relations are shown as arrows. The domain model used to structure MWCC content has been developed in close collaboration with experts from the NOAA Climate Program Office (see Section 5).

At the time of writing, the authors are in the process of transforming the bitmap-based tag cloud and ontology graph components to dynamic *Scalable Vector Graphics* (SVG) representations, supporting on-the-fly computations instead of weekly updates. This will enable us to extend the synchronization mechanism and highlight terms related to the currently edited document in the tag cloud, and select the ontology concept that best describes the content of the edited document. Topical shifts by the co-authors, therefore, will translate into a panning operation and the display of different, more relevant concepts.

5. Implementation and Evaluation

Following an evolutionary and user-centered development approach, rapid feedback cycles have been instrumental in MWCC conceptualization and implementation. The evaluation activities focused on usability, the accuracy of extracted knowledge, and the added value in real-world use cases.

Initially we conducted *usability inspections*, asking colleagues and experts from partner organizations to assess the interface design against recognized usability principles. More comprehensive *usability testing* in later phases of the project gathered feedback from actual users, including the employees of large organizations such as the *NOAA Climate Program Office* (www.climate.gov), he *National Cancer Institute* (www.cancer.gov), and the *Austrian Chamber of Commerce* (www.wko.at). The explicit feedback from the usability testing was combined with implicit feedback from monitoring user activities – in 2012, for example, about 4,600 users performed more than 13,000 searches.

MWCC’s evolutionary development approach supports iterative, user-driven refinements and the integration of new features early in the development cycle. The NOAA Climate Program Office, for example, has been using the analytical MWCC components for several years and suggested numerous improvements. NOAA employees often engage in public outreach activities and collaboratively author articles for public dissemination. The MWCC editing environment supports this task with real-time feedback on how other stakeholders perceive topics of interest. While authors type, the system analyzes the evolving document and recommends related articles from the repositories of NOAA, or from external sources. At any time during the editing process, authors can inspect these resources, for example to minimize overlaps or include references. Showing positive and negative associations with a given topic helps authors to choose engaging, positive terms and stay clear of those with a negative connotation. Detecting such associations and suggesting related content items are key elements of the MWCC editing environment. Therefore the evaluation focused not only on the usability of the system as outlined above, but also used *Climate Quiz* [10], a crowdsourcing application in the tradition of games with a purpose, to assess the quality of extracted relations.

Two workshops held in June 2010 and September 2012 helped to gather feedback from international climate change experts and environmental stakeholders from various sectors including research centers, non-profit organizations, companies, and government agencies. The participants pointed out the merging of the traditionally disparate processes of information seeking and authoring as the main benefit when working on press releases and joint statements. They also requested additional export functions as outlined in Section 3 to further process the collected data in third-party applications. Additional feedback was gathered through presentations of the prototype at side events of the *UN Conference on Sustainable Development* (Rio+20) and the *UN Conference on Climate Change* (COP-18), as well as the *Annual Meeting of the American Association for the Advancement of Science* (AAAS-2013).

6. Summary and Conclusion

Environmental stakeholders including policy advisors, public outreach departments and campaign organizers have recognized that the Internet is not just a medium to obtain information, but a catalyst of ad-hoc communication and collaboration processes between individuals who want to share their opinions and jointly create knowledge resources. By uncovering patterns and trends, the *webLyzard.com* platform and its portfolio of semantic technologies increases the transparency of these processes and replaces static repositories by an agile, collaborative framework to manage evolving knowledge.

The *Media Watch on Climate Change* (MWCC) is a domain-specific portal built upon this platform to support environmental decision makers with consolidated climate change knowledge from multiple sources. It uses automated methods to annotate and classify documents while they are being edited, and to recommend related content. Interactive visualizations serve as navigational aids and reflect the semantic and geospatial context of the edited document, as well as topical shifts introduced by the authors. By providing the portal as a public resource, the authors hope to promote the adoption of semantic technologies, and to close the gap that exists “between users’ desire to collaborate and the capabilities of the tools they use” [7]. The system provides a range of collaborative features, including the concurrent editing of documents, the joint definition of topics, and the export of shared resources for reuse in external applications – supporting the existing workflows of environmental stakeholders.

Acknowledgements. The collaborative editing components of the Web portal we present were developed within DIVINE (www.weblyzard.com/divine) and Triple-C (www.ecoresearch.net/triple-c),two research projects funded within the FIT-IT Semantic Systems program line of the Federal Ministry for Transport, Innovation, and Technology (BMVIT) managed by the Austrian Research Promotion Agency (FFG), and the ACRP Program of the Austrian Climate and Energy Fund.

7. References

1. Ahn, L.v. (2006). “Games with a Purpose”, Computer, 39(6): 92-94.
2. Breslin, J.G. and Decker, S. (2007). “The Future of Social Networks on the Internet: The Need for Semantics”, IEEE Internet Computing, 11(6): 86-90.
3. Chen, H. (2010). “Business and Market Intelligence 2.0”, IEEE Intelligent Systems, 25(1): 68-83.
4. Doan, A., Ramakrishnan, R. and Halevy, A.Y. (2011). “Crowdsourcing Systems on the World Wide Web”, Communications of the ACM, 54(4): 86-96.
5. Finholt, T.A. (2002). “Collaboratories”, Annual Review of Information Science and Technology, 36(1): 73-107.

6. Fischer, G. (2011). “Understanding, Fostering, and Supporting Cultures of Participation”, interactions, 18(3): 42-53.

7. Golovchinsky, G., Qvarfordt, P. and Pickens, J. (2009). “Collaborative Information Seeking”, Computer, 42(3): 47-51.

8. Hubmann-Haidvogel, A., Scharl, A. and Weichselbraun, A. (2009). “Multiple Coordinated Views for Searching and Navigating Web Content Repositories”, Information Sciences, 179(12): 1813-1821.

9. Scharl, A., Hubmann-Haidvogel, A., et al. (2013). Media Watch on Climate Change – Visual Analytics for Aggregating and Managing Environmental Knowledge from Online Sources. 46th Hawaii International Conference on Systems Sciences (HICSS-46). R.H. Sprague. Maui, USA: IEEE Press: 955-964.

10. Scharl, A., Sabou, M. and Föls, M. (2012). Climate Quiz – A Web Application for Eliciting and Validating Knowledge from Social Networks. 18th Brazilian Symposium on Multimedia and the Web (WebMedia-2012). G. Bressan and R.M. Silveira. São Paulo, Brazil: ACM. 189-192.

11. Scharl, A. and Tochtermann, K., Eds. (2007). The Geospatial Web - How Geo-Browsers, Social Software and the Web 2.0 are Shaping the Network Society. London: Springer.

12. Simperl, E., Thurlow, I., et al. (2010). “Overcoming Information Overload in the Enterprise - The Active Approach”, IEEE Internet Computing, 14(6): 39-46.

13. Syed, K.A.A., Kröll, M., et al. (2012). “Incremental and Scalable Computation of Dynamic Topography Information Landscapes”, Journal of Multimedia Processing and Technologies 3(1): 49-65.

14. Weichselbraun, A., Gindl, S. and Scharl, A. (2013). “Extracting and Grounding Context-Aware Sentiment Lexicons”, IEEE Intelligent Systems: Forthcoming (Accepted 06 Jan 2013).