Embedded Human Computation for Knowledge Extraction and Evaluation (uComp)

D6.3: Scientific Impact, Dissemination, Exploitation Planning and Community Engagement Report

22 August 2016
Version: 1.0

Version History

Version	Date	Author	Comments
0.1	2016-06-20	M Lang	Initial Version and Document Structure
0.1x	2016-06-31	Consortium	Partner Contributions
0.2	2016-08-08	A Scharl	Revision and Extension
0.3	2016-08-16	M Lang	Major Revision
0.31	2016-08-17	A Scharl	Minor Revision
1.0	2016-08-22	A Scharl	Final Version

Dissemination Level: PU – Public

Table of Contents

Overview

This report outlines the dissemination and impact activities that have been been performed in the uComp project, what future activities are planned, and how the project results will be made available to the European industry. Furthermore it outlines exploitation activities undertaken and the planned exploitation strategies to maximise outcome after the project.

All partners will continue to play active roles in dissemination, placing special emphasis upon scientific channels of publications including journal and conference papers, but also actively engaging in workshop organization, courses, invited talks, and various online dissemination activities. External networks and advisors form a key part of the dissemination activities towards non-scientific users, as they are providing outreach towards citizens and organizations in diverse sectors.

The main public interface of uComp is the project web site (www.ucomp.eu), which contains information on the project objectives, partners, R&D, a short presentation and results achieved. The scientific results of the project have been submitted to major international conferences and journals. The project has taken part in relevant national and European concertation events and plans to continue to do so.

In order to promote maximum use and dissemination, major technology components from uComp are being made available as open source and are thus easily exploitable both commercially and for research purposes beyond the end of the project. In order to ensure support beyond the project's lifetime, the source code of these components is and will remain available on well-known servers (e.g., GitHub).

Dissemination and Community Engagement

For maximum impact, uComp has adopted a multi-channel dissemination approach outlined in the following sections. We put special emphasis on the project's Web site and Twitter presence, and used social media in conjunction with two human computation applications in the tradition of games with a purpose to engage players and increase the visibility of uComp within the scientific community.

Web Site

The project has set up and maintains a public Web site at www.ucomp.eu, based on the WordPress content management system to enable distributed editing among consortium members. The site functions as a dissemination and documentation tool, and single access point that references uComp applications and services. In addition to documenting the scientific output of the project, examples and descriptions outline how uComp information services can be used by third parties, independently or in conjunction with other Web applications. The project partners will ensure that the website will remain available for at least three years after the end of the project.
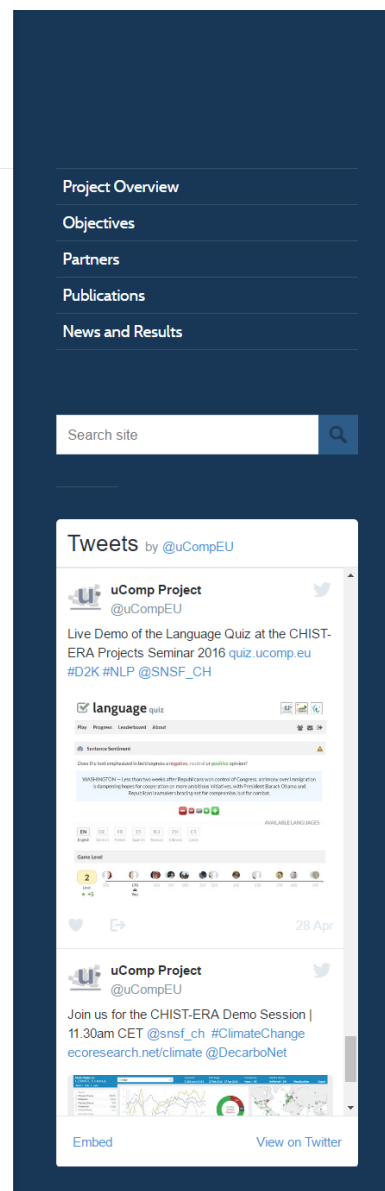
Figure 1. Screenshot of the uComp.eu Web site

Social Media

Posting project-related information to social media platforms extends outreach and increases the exposure of project results. The uComp project has established a Twitter account (@uCompEU) early on in the project, giving the project partners a channel to disseminate their publications, events organised and participated in, etc.

As of August 2016, the Twitter account has 156 followers and published 124 tweets. Project members also disseminate uComp-specific results in additional social media channels such as LinkedIn and Facebook. Either the account name @uCompEU or the hashtag #uCompEU is used to tag posts published via these channels.

Figure 2. Screenshot of the project's Twitter account @uCompEU

Impact Metrics

The following diagrams reflect the impact of the project's social media activities via the @uCompEU Twitter account, including the Top Tweets in the first two quarters of 2016 between April and June 2016 (including the CHIST-ERA Projects Seminar held in Bern, Switzerland). In terms of audience statistics, the female-to-male ratio was 31:69, and the top countries were the United Kingdom (19%), Austria (12%), the United States (12%), Italy (8%), Spain (8%) and France (5%).
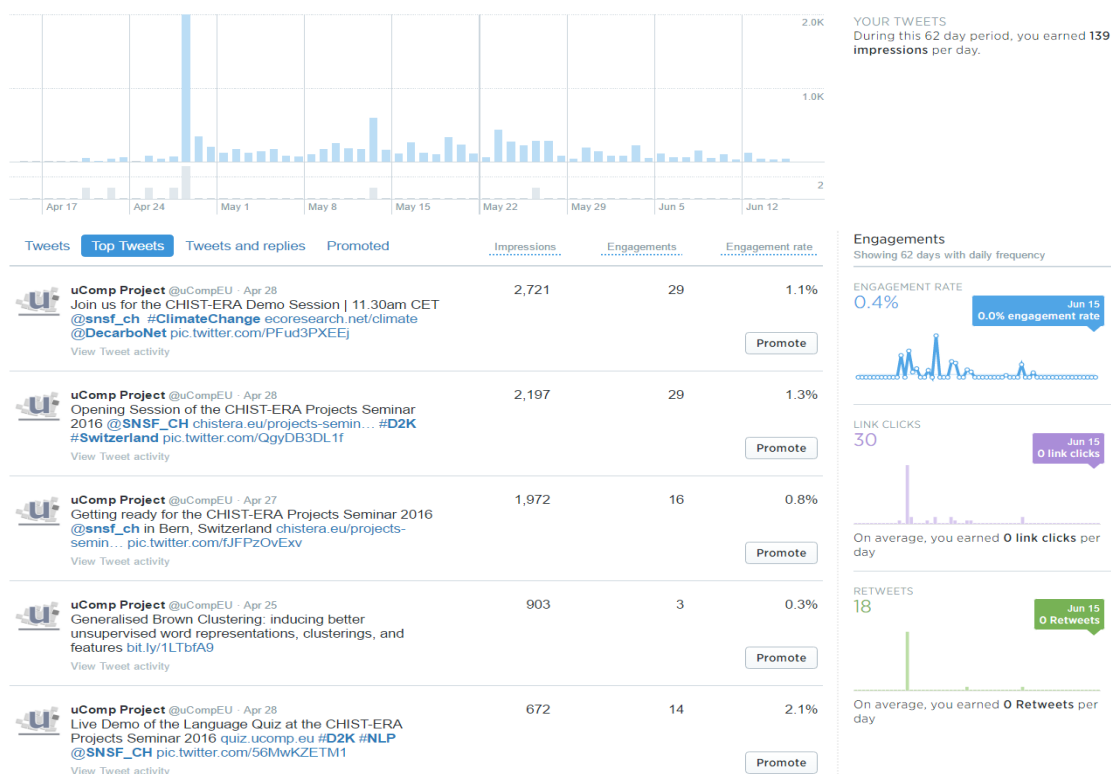
Figure 3. Impact of @uCompEU's Twitter activities

Climate Challenge

The *Climate Challenge* is an online competition in the tradition of games with a purpose that combines practical steps to reduce carbon footprint with predictive tasks to estimate future climate-related conditions. The application is designed to increase environmental literacy and motivate users to adopt more sustainable lifestyles. Its feedback channels include a leaderboard and a visual tool to compare answers of individual players with (i) the average assessments of their direct social network contacts as well as the entire pool of participants, (ii) a selected group of experts, and (iii) real-world observations (Scharl et al., 2015).

Since the main goal of the Climate Challenge is not only to collect answers, but also to raise awareness, the promotion and community aspect of the game has to be treated different than in the Language Quiz (see following section). There are three different categories of crowdsourcing paradigms, and it is hard to mix the monetary reward system of CrowdFlower with a gameplay-driven and altruistic reward system of the Climate Challenge. Instead, social media channels such as Facebook and Twitter were used to promote the game and create a stable community.
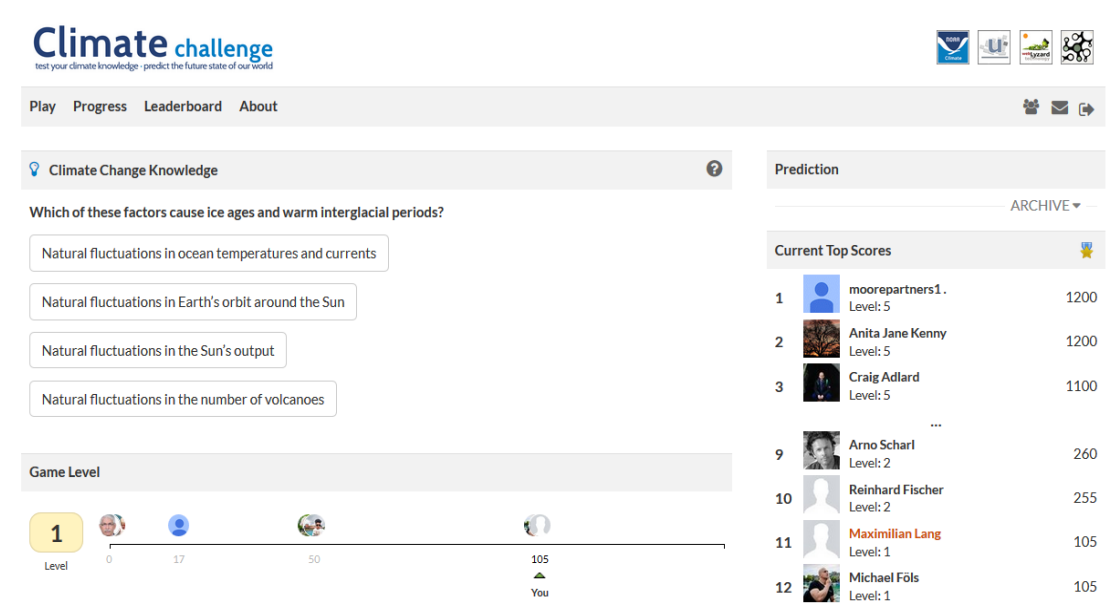
Figure 4. Main interface elements of the Climate Challenge[1]

Another special aspect of the Climate Challenge is the recurring nature of the tasks. Inevitably, there is is a period where a prediction task is "open" for answers, followed by a waiting period until the real-world answer becomes available. Then the next prediction question can be asked. To account for this delay, and as an incentive

---

[1] www.ecoresearch.net/climate-challenge

measure, monthly game rounds were introduced where only a certain number of tasks per task type are being made available. This approach ensures that players will always find new tasks at the beginning of each month, and have a incentive to return to the game to find out about the correct answer to the previous question.

The created community was used as a platform to promote the new round each month, supplemented by e-mail notifications and Facebook announcements (to the approximately 2,500 followers of the Facebook community page) that new results were available. This approach was used as a regular reminder to re-activate players and keep them engaged each month.

Since its launch in March 2015, the Climate Challenge attracted 4,379 unique users as of August 2016. From those, 873 created a user-account in the game leading to a high conversion rate of 19.93%.

Out of those 873 users, 738 became active players. In total the game collected 722 prediction answers, 5,973 multiple choice answers, 23,380 sentiment answers, 2,802 pledge answers and 224 answers to the opinion poll.

Language Quiz

The approach of the *Language Quiz* differs significantly from the *Climate Challenge*, even though both games are built upon the same framework. Instead of creating environmental awareness and promoting sustainable lifestyle choices, the Language Quiz aims to acquire multi-lingual language resources for research purposes.

A flexible and open architecture to support multiple languages, different types of tasks, and the inclusion of third-party tasks were among the key factors that have guided the development efforts. Conceptual insights gathered from analysing the results of the *Sentiment Quiz*, a game with a purpose to assess sentiment terms (Scharl et al., 2012), also influenced the design of the *Language Quiz* - although the two applications are based on a completely different technology stack.

Since its launch in October 2015, the Language Quiz attracted 3,469 unique users. 2,147 of these users created a user account, and 1,999 became active players. Out of the 1,999 active players, 1,916 were contacted via the CrowdFlower campaign option. The game was also promoted at RANLP-2015 and through social media.

956 users became active players who submitted valid answers, 1,043 failed the test questions. In total 64,961 valid answers were submitted to the game, 55,791 of those answers where sourced via CrowdFlower, and 9,170 where organic answers from players recruited via other channels.

Figure 5. Main interface elements of the Language Quiz[2]

Scientific Publications and Related Public Outreach Activities

As documented on the uComp web site, we have published and are publishing a growing body of scientific papers and presentations at technical and scientific conferences. Specific outlets were selected based on their relation to the topics addressed by uComp, and on their impact. The publications are aimed at sharing the results of uComp with the European and international scientific community, to encourage their incorporation into the work of other scientists and technologists. The project also provides online software demonstrations, available from the uComp Website.

With respect to intellectual property rights and promoting take-up of project results, the consortium partners make all scientific results of this project available to the research community with links to open access articles where possible. In addition, associated tools are released in the form of open source software via GitHub.com to facilitate take-up.

In order to ensure the widest possible audience, uComp has used press releases aimed at the academic and industrial sectors (e.g., delivery of tutorials), industry-oriented dedicated workshops, etc. (some example of newsletters, media coverage, activities, etc.).

---

[2] quiz.ucomp.eu

Refereed Scientific Publications

The scientific output of the uComp project includes 35 refereed scientific publications; two additional papers have been submitted and are currently under review.

2016

- Derczynski, L., Bontcheva, K., Roberts, I. (2016). Broad Twitter Corpus: A Diverse Named Entity Recognition Resource. Submitted to COLING 2016..

- Derczynski, L. (2016). Complementarity, F-score, and NLP Evaluation. In Proceedings of LREC, Forthcoming.

- Derczynski, L., Strötgen, J., Maynard, D., Greenwood, M.A., Jung, M. (2016). GATE-Time: Extraction of Temporal Expressions and Events. In Proceedings of LREC, Forthcoming.

- Derczynski, L., Chester, S. (2016). Generalised Brown Clustering and Roll-Up Feature Generation. In Proceedings of AAAI 2016, Forthcoming.

- Paroubek, P. (2016). Critères pour l'annotation active de microblogs. In Actes de l'atelier Communautés en ligne : outils et applications TAL (COLTAL), JEP-TALN-RECITAL Conference.

- Scharl, A., Föls, M., Herring, D., Piccolo, L., Fernandez, M., Alani, H. (2016). Climate Challenge – A Game with a Purpose to Promote Collective Awareness and Behavioural Change. 3rd International Conference on Internet Science (INSCI-2016), Forthcoming.

- Scharl, A., Hubmann-Haidvogel, A., Jones, A., Fischl, D., Kamolov, R., Weichselbraun, A., Rafelsberger, W. (2016). Analyzing the public discourse on works of fiction – Detection and visualization of emotion in online coverag eabout HBO's Game of Thrones. Information Processing & Management, Volume 52, Issue 1, pp. 129-138.

- Vakulenko, S., Weichselbraun, A., Scharl, A. (2016). Detection of Valid Sentiment-Target Pairs in Online Product Reviews and News Media Articles. In Proceedings of WI 2016 2016 IEEE/WIC/ACM International Conference on Web Intelligence, Forthcoming.

- Weichselbraun, A., Scharl, A.; Gindl, S. Extracting Opinion Targets from Environmental Web Coverage and Social Media Streams. In Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS-49), pp. 1040-1048.

- Wohlgenannt, G., Sabou, M., Hanika, F. (2016). Crowd-based Ontology Engineering with the uComp Protege Plugin. Semantic Web Journal 7 (4), pp. 379-398

- Wohlgenannt, G., Minic, F. (2016). Using word2vec to Build a Simple Ontology Learning System. In Proceedings of the ISWC 2016 Posters and Demonstrations Track, 15th International Semantic Web Conference, ISWC 2016, Forthcoming.

2015

- Belk, S., Wohlgenannt, G., Polleres, A. (2015). Exploring and exploiting(?) the awkward connections between SKOS and OWL. Proceedings of the ISWC 2015 Posters & Demo. Track, In CEUR Workshop Proceedings, Vol. 1486.

- Derczynski, L., Maynard, D., Rizzo, G., van Erp, M., Gorrell, G., Troncy, R., Petrak, J., Bontcheva, K. (2015). Analysis of named entity recognition and linking for tweets. Information Processing & Management 51(2), pp. 32–49.

- Derczynski, L., Augenstein, I., Bontcheva, K. (2015). USFD: Twitter NER with Drift Compensation and Linked Data, In Proceedings of the Workshop on Noisy User-Generated Text (W-NUT).

- Derczynski, L., Bontcheva, K. (2015). Efficient Named Entity Annotation through Pre-empting, In Proceedings of Recent Advances in Natural Language Processing, ACL.

- Fraisse, A., Hamon, T., Grouin, C., Paroubek, P., Zweigenbaum, P. (2015). Proceedings of the DEFT 2015 Eleventh Text Mining Challenge (DEFT 2015 Onzième Défi Fouille de Texte), workshop at TALN 2015 the 22nd Conference on Natural Language Processing (TALN 2015).

- Fraisse A., Paroubek, P. (2015). Vers des pratiques collaboratives pour les systèmes d'organisation de connaissances (Towards collaborative Practices for Knowledge Management Systems). In Proceedings of the Tenth Conference of ISKO-France 2015, p 16.

- Scharl, A., Föls, M., Herring, D. (2015). Climate Challenge – Raising Collective Awareness in the Tradition of Games with a Purpose. 14th Brazilian Symposium on Human Factors in Computer Systems (IHC-2015).

- Wohlgenannt, G. (2015). Heterogeneous Sources of Evidence in Ontology Learning. In Proceedings of ESWC 2015, Vol. 9088, pp. 54-68.

- Wohlgenannt, G., Belk, S., Rohrer, K. (2015). Optimizing Ontology Learning Systems that use Heterogeneous Sources of Evidence. In MIWAI 2015, pp. 137-148.

2014

- Bontcheva, K., Roberts, I., Derczynski, L., Rout, D. (2014). The GATE Crowdsourcing Plugin: Crowdsourcing Annotated Corpora Made Easy. In Proceedings of the European Association for Computation Linguistics, 2014.

- Bontcheva, K., Derczynski, L., Roberts, I. (2014). Crowdsourcing Named Entity Recognition and Entity Linking Corpora. In Ide, N. and Pustejovsky,, J. (eds), The Handbook of Linguistic Annotation.

- Derczynski, L., Maynard, D., Rizzo, G., van Erp, M., Gorrell, G., Troncy, R., Petrak, J., Bontcheva, K (2014). Analysis of Named Entity Recognition and Linking for Tweets. Information Processing & Management, 51(2), pp. 32-49.

- Fischl, D., Scharl, A. (2014). Metadata Enriched Visualization of Keywords in Context, Sixth ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS-2014), pp. 193-196.

- Fraisse, A., Paroubek, P. (2014). Toward a Unifying Model for Opinion, Sentiment and Emotion Annotation and Information Extraction. 9th Language Resources and Evaluation Conference (LREC-2014).

- Fraisse A., Paroubek, P. (2014). Twitter as a Comparable Corpus to build Multilingual Affective Lexicons. In Proceedings of the 7th International Workshop on Building and Using Comparable Corpora at LREC 2014 (BUCC 2014).

- Hanika, F., Wohlgenannt, G., Sabou, M. (2014).The uComp Protege Plugin: Crowdsourcing Enabled Ontology Engineering. Knowledge Engineering and Knowledge Management – 19th International Conference, EKAW 2014, pp. 181-196.

- Hanika, F., Wohlgenannt, G., Sabou, M. (2014). The uComp Protege Plugin for Crowdsourcing Ontology Validation. In Proceedings of the ISWC 2014 Posters and Demonstrations Track a track within the 13th International Semantic Web Conference, ISWC 2014, pp. 253-256.

- Pak, A., Paroubek, P., Fraisse, A., Francopoulo, G. (2014). Normalization of Term Weighting Scheme for Sentiment Analysis. In the Springer Verlag series: (LNAI) Lecture Notes in Artificial Intelligence.

- Sabou, M., Bontcheva, K., Derczynski, L., Scharl, A. (2014). Corpus Annotation through Crowdsourcing: Towards Best Practice Guidelines. 9th Language Resources and Evaluation Conference (LREC-2014).

- Scharl, A., Kamolov, R., Fischl, D., Rafelsberger, W., Jones, A. (2014). Visualizing Contextual Information in Aggregated Web Content Repositories. 9th Latin American Web Congress (LA-WEB 2014). J.M. Almeida et al. Ouro Preto, Brazil: IEEE Press: 114-118.

- Scharl, A., Hubmann-Haidvogel, A., Rafelsberger, W., Weichselbraun, A., Lang, H.P., Sabou, M. (2014). Visualizing Knowledge Along Semantic and Geographic Dimensions – A Web Intelligence Platform to Explore Climate Change Coverage. In Okada, A., Buckingham Shum, S., Sherborne, T. (eds), Knowledge Cartography: Software Tools and Mapping Techniques, pp. 421-439.

- Weichselbraun, A., Gindl, S., Scharl, A. (2014). Enriching Semantic Knowledge Bases for Opinion Mining in Big Data Applications, Knowledge-Based Systems, 69, pp. 78-85.

- Wohlgenannt, G., Belk, S., Karacsonyi, M., Schett, M. (2014). Using an Ontology Learning System for Trend Analysis and Detection. In Proceedings of the ISWC 2014 Posters and Demonstrations Track a track within the 13th International Semantic Web Conference, ISWC 2014, pp. 37-40.

2013

- Bontcheva, K., Derczynski, L., Funk, A., Greenwood, M.A., Maynard, D., Aswani. N. (2013). TwitIE: An Open-Source Information Extraction Pipeline for Microblog Text. Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2013). Download TwitIE.

- Derczynski, L., Maynard, D., Aswani, N., Bontcheva, K. (2013). Microblog-Genre Noise and Impact on Semantic Annotation Accuracy. 24th ACM Conference on Hypertext and Social Media.

- Derczynski, .L., Yang, B., Jensen, C.S. (2013). Towards Context-Aware Search and Analysis on Social Media Data. In Proceedings of the Extending Database Technology conference (EDBT 2013).

- Derczynski, L., Ritter, A., Clark, S., Bontcheva, K. (2013). Twitter Part-of-Speech Tagging for All: Overcoming Sparse and Noisy Data. In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2013). Download the PoS tagger.

- Fraisse, A., Paroubek, P., Francopoulo, G. (2013). Improving Minor Opinion Polarity Classification with Named Entity Analysis. 20th Conference on Natural Language Processing (Traitement Automatique du Langage Naturel) (TALN 2013).

- Gindl, S., Weichselbraun, A., Scharl, A. (2013). Rule-based Opinion Target and Aspect Extraction to Acquire Affective Knowledge. WWW Workshop on Multidisciplinary Approaches to Big Social Data Analysis (MABSDA-2013, pp. 557-563.

- Sabou, M., Scharl, A., Föls, M. (2013). Crowdsourced Knowledge Acquisition: Towards Hybrid-Genre Workflows, International Journal on Semantic Web and Information Systems, 9(3), pp. 14-41.

- Sabou, M., Bontcheva, K., Scharl, A., Föls, M. (2013). Games with a Purpose or Mechanised Labour? A Comparative Study. In Proceedings of 13th International Conference on Knowledge Management and Knowledge Technologies (i-KNOW)

- Scharl, A., Herring, D. (2013). Extracting Knowledge from the Web and Social Media for Progress Monitoring in Public Outreach and Science Communication. 19th Brazilian Symposium on Multimedia and the Web (WebMedia-2013), pp.121-124.

- Weichselbraun, A., Scharl, A., Lang, H.-P. (2013). Knowledge Capture from Multiple Online Sources with the Extensible Web Retrieval Toolkit (eWRT). Seventh International Conference on Knowledge Capture (K-CAP 2013).

- Wohlgenannt, G, Belk, S., Schett, M. (2013). Computing Semantic Association: Comparing Spreading Activation and Spectral Association for Ontology Learning. In Computing Semantic Association: Comparing Spreading Activation and Spectral Association for Ontology Learning, pp. 317-328.

- Wohlgenannt, G., Belk, S., Schett, M. (2013). A Prototype for Automating Ontology Learning and Ontology Evolution. In 5th International Conference on Knowledge Engineering and Ontology Development (KEOD-2013).

Press Releases and Media Coverage

Due to the various dissemination activities such as press releases, the project has sparked notable media interest resulting in international media attention and coverage throughout the project. The list below lists the various items of media coverage, copies of the press releases along with a selection of the media coverage items can be found in the appendix.

Press Releases

Two major press releases have been issued, the first one by MODUL in June 2013 and the second one by NOAA in collaboration with MODUL to introduce the Climate Challenge Game with a purpose in March 2015.

- **National Oceanic and Atmospheric Administration (NOAA)**, 24/03/2015, "NOAA unveils Climate Challenge: an interactive game to build and test climate knowledge",
  http://cpo.noaa.gov/OutreachandEducation/OutreachArchive/TabId/505/ArtMID/1254/ArticleID/242371/NOAA-unveils-Climate-Challenge-an-interactive-game-to-build-and-test-climate-knowledge.aspx

- **MODUL University Vienna**, 26/06/2013, "Opinion Lead versus Information Overload: Automatic – But Smart – Use of Social Media",
  https://www.modul.ac.at/article/view/opinion-lead-versus-information-overload-automatic-but-smart-use-of-social-media/

Media Coverage

2015

- **Bluewin.ch**, 10/08/2015, "Online-Spiel «Climate Challenge» unterstützt Klimaschutz",
  http://www.bluewin.ch/de/leben/nachhaltigkeitsblog/2015/15-08/wir-schweizer-sind-umweltbewusster-geworden.html

- **IT&T Business**, 03/04/2015, "Schwarmintelligenz gegen Expertenwissen", p.8

- **Umweltdialog.de**, 30/03/2015, "Klimawandel wird mit Social-Media-Spiel bekämpft",
  http://www.umweltdialog.de/de/umwelt/klima/2015/Klimawandel-wird-mit-Social-Media-Spiel-bekaempft-.php

- **Die Presse**, 28/03/2015, "Klima: Forscher wollen mit Spiel sensibilisieren", p.36

- **ICT Kommunikation**, 27/03/2015, "Mit Social Media den Klimawandel bekämpfen",
  http://ictk.ch/content/mit-social-media-den-klimawandel-bek%C3%A4mpfen

- **Krone.at**, 27/03/2016, "Wiener Game soll für Klimawandel sensibilisieren",
  http://www.krone.at/Digital/Wiener_Game_soll_fuer_Klimawandel_sensibilisieren-Climate_Challenge-Story-445588

- **Enertipp.ch**, 26/03/2015, "Forscher wollen Menschen mit kompetitivem Game sensibilisieren",
  http://www.enertipp.ch/?z=410&id=2685

- **Greenpeace Magazin**, 27/03/2015, "Spielen gegen den Klimawandel",
  https://www.greenpeace-magazin.de/nachrichtenarchiv/spielen-gegen-den-klimawandel

- **Pressetext.com**, 26/03/2015, "Klimawandel wird mit Social-Media-Spiel bekämpft",
  http://www.pressetext.com/news/20150326019

2014

- **FWF Info**, 24/10/2014, "Am Puls der Wissenschaft",
  https://www.fwf.ac.at/fileadmin/files/Dokumente/info-Magazin/info89-14-02.pdf

- **WCM**, 7/02/2014, "Soziale Medien automatisch aber intelligent nutzen", p. 38

2013

- **MAS Magazin**, 20/09/2013, "Wissensvorsprung statt Datenflut: Soziale Medien automatisch – aber intelligent – nutzen", p.6

- **IX Magazin**, 01/08/2013, "Wissen aus sozialen Netzwerken ziehen", p. 15

- Information-age.com, 09/07/2013, "Using man and machine to understand the web",
  http://www.information-age.com/technology/information-management/123457187/using-man-and-machine-to-understand-the-web--

- **Telepolis**, 06/07/2013, "ProfessorIn für Data Mining gesucht",
  http://www.heise.de/tp/artikel/39/39420/1.html

- **IT Daily**, 05/07/2013, "Wissensvorsprung statt Datenflut",
  http://www.it-daily.net/it-technologie/aktuelles/7695-wissensvorsprung-statt-datenflut

- **Medianet**, 05/07/2013, "Social-Media Daten automatisiert auswerten",
  http://medianet.at/front/download_singlepage/5547

- **cmo.com.au**, 05/07/2013, "University brings human intellect and technology together to solve social media puzzle", http://www.cmo.com.au/article/472047/university_brings_human_intellect_technology_together_solve_social_media_puzzle/

- **Bieler Tagblatt**, 26/06/2013, **"**Erfolge im Forschungs-Projekt uComp der MODUL University Vienna", http://www.bielertagblatt.ch/unterhaltung/digital/erfolge-im-forschungs-projekt-ucomp-der-modul-university-vienna

- **phys.org**, 26/06/2013, "uComp research project delivers first results under open source license", http://phys.org/news/2013-06-ucomp-results-source.html

- **relevant.at**, 26/06/2013, "Opinion Lead versus Information Overload: Automatic – But Smart – Use of Social Media", http://relevant.at/wirtschaft/print/1004186/opinion-lead-versus-information-overload-automatic-but-smart-use-of-social-media.story

- **Radio Ö1 (audio segment)**, 19/06/2013, "digital.leben: Zwischen Schandlohn und Chance Crowdsourcing in der Wissenschaft", http://oe1.orf.at/programm/340723

Other Dissemination Results and Activities

2016
- The 9th GATE Training Course: Mining social media content with GATE (includes a module on crowdsourcing for text mining). June 6th – 10th 2016, Sheffield, UK;

2015

- FOI, the Swedish Defence Research agency, started using our TwitIE toolkit in detecting information warfare (e.g. Russian and European social media interactions around Ukraine);

- The 8th GATE Training Course: Mining social media content with GATE (includes a module on crowdsourcing for text mining). June 8th – 12th June 2015, Sheffield, UK;

- uComp provided the gold standard data of DEFT 2015 evaluation campaign about opinion mining of Tweets about Climate Change, whose results were presented at the DEFT 2015 workshop at TALN 2015 the 22nd Conference on Natural Language Processing (TALN 2015 22ème Conférence sur le Traitement Automatique des Langues), June 22nd 2015, Caen, France;

- Launch of the Climate Challenge (March 2015);
  http://www.ecoresearch.net/climate-challenge

- Launch of the Language Quiz (June 2015);
  http://quiz.ucomp.eu

2014

- A half-day tutorial on NLP for social media and crowdsourcing was delivered at EACL'2014 in Sweden.

- A course on how humans use language across social networks, involving multi-network effects and linguistic analysis, containing also uComp material was delivered in Tübingen (August 2014) as part of ESSLLI.

- The 7th GATE Training Course: Mining social media content with GATE (included a module on crowdsourcing for text mining) June 9th – 13th 2014 in Sheffield, UK.

- Tutorial: Natural Language Processing for Social Media. K. Bontcheva and L. Derczynski on April 26th 2014 at EACL 2014.

2013

- The 6th GATE Training Course was held in Sheffield, UK on June 3 - 7 2013 which had a dedicated module on mining social media.

- A tutorial on NLP for social media and crowdsourcing will be delivered as a one week course to PhD students in Hungary (December 9 - 13 2013).

Project-Wide Achievements and Exploitation Strategies

The uComp partners plan to exploit the methodological advances of uComp individually and in collaboration. Due to the generic nature of the results, exploitation activities will go beyond specific industries and defined uses cases, as outlined in the following sections. They will center on a number of key components including (i) the HC framework including GATE and Protégé plugins and REST API, (ii) the semantic knowledge base, and (iii) improved extraction components for factual and affective knowledge.

University of Sheffield (USFD)

The project enabled USFD to build a strong reputation in crowdsourcing and human computation for natural language processing and knowledge extraction. An open-source plugin with reusable tools and algorithms has been released, as a result and

the corresponding EACL'2014 paper has already received 12 citations. As the software tools are openly available on SourceForge, as part of the GATE NLP framework, it is not possible; unfortunately, to pinpoint the exact number of downloads for the crowdsourcing tools themselves. GATE itself has been downloaded 102,855 times since 1 Jan 2014 to 13 August 2016, with the US as the top country (20%), followed by India, United Arab Emirates, Germany, and the UK.

We have also publicised the uComp findings and results during our annual training course on GATE for social media analytics.

Another key result, which has only just been released, is the biggest openly available corpus of crowdsourced and adjudicated tweets annotated with named entities (Person, Location, Organisation). A paper describing the corpus and its properties, as well as comparing it to other similar corpora, is currently under review for COLING'2016. A pre-release of the work was included with a popular article published in Information Processing and Management journal in 2015.

Expertise and tools developed in uComp originally, were then taken up and used in the PHEME project, to create a dataset of crowdsourced rumourous conversations:

Crowdsourcing the annotation of rumourous conversations in social media, A Zubiaga, A., Liakata, M., Procter, R., Bontcheva, K., Tolmie, P. (2015). In Proceedings of the 24th International Conference on World Wide Web, pp. 347-353.

USFD published over 14 papers in total, with the 6 key papers on crowdsourcing and human computation receiving so far over 100 citations. The other set of very highly cited papers revolve around the topic of natural language processing for noisy data, especially named entity recognition. The 8 key papers here have over 285 citations, with the two most highly cited papers from 2013 receiving 87 and 85 citations respectively. There are two more papers under review (COLING'2016 and IEEE Computer Speech and Language) and two more papers in preparation.

The GATE crowdsourcing plugin has now been taken up for use, maintenance, and, if needed, further development as part of the following research projects:

- SoBigData - an H2020 research infrastructure project on social media analytics, where human computation will play an important role;

- DecarboNet - an FP7 project, which is using the uComp GATE crowdsourcing plugin to create gold-standard datasets for evaluation of entity linking, environmental term extraction, and opinion mining;

- COMRADES - an H2020 CAPS project, which will use the uComp GATE crowdsourcing plugin to create much needed gold-standard datasets for evaluating algorithms for analysing disaster response social media content.

USFD, as a research-focused institution, will exploit the knowledge and results of uComp as grounds for new scientific publications; creation of gold-standard datasets; new research projects; and scientific collaborations. USFD will also continue teaching crowdsourcing and human computation at the annual GATE training course, and also release openly available training modules as part of the SoBigData training activities.

The open-source GATE crowdsourcing plugin will continue to be maintained and supported, which forms another promising avenue for exploitation through GATE consulting activities.

MODUL University Vienna (MOD)

The methodological insights and visibility gained from the uComp project helped to underscore MOD's leading expertise in GWAP development. The project not only helped to advance the *Media Watch on Climate Change*, but also strengthened the collaboration with the National Oceanic and Atmospheric Administration (NOAA), and helped to establish a project with the *United Nations Environment Programme* (UNEP) as another international organisation that is using the algorithms of MOD and WU to extract factual and affective knowledge.

These knowledge extraction algorithms are at the core of MOD's research activities; the hybrid uComp framework to gather and validate the required language resources has already been very useful in this context, and will continuously be leveraged in ongoing and future research initiatives - e.g. *InVID - In Video Veritas*, a three-year H2020 Innovation Action (Start: January 2016) that builds a knowledge verification platform to detect emerging stories and assess the reliability of newsworthy video files and content spread via social media.

The uComp human computation framework has also become an important part of the MOD curriculum - the developed applications (Language Quiz, Climate Challenge) are used in Bachelor and Master programs to explain the fundamentals of crowdsourcing and games with a purpose, and the role of incentive mechanisms and engagement strategies to promote the uptake of social media application. The work conducted in uComp is taught in the MBA program of MOD as part of an annual core unit on *Social Media Intelligence*. Within the PhD program of MOD, the achieved progress is conveyed as part of a unit on *Knowledge Extraction and Verification* - both in terms of knowledge extraction algorithms, but also in terms of using human computation to acquire gold standard data.

MOD spin-outs (MODUL Research, MODUL Technology, webLyzard) will directly benefit from the projects' research output by diversifying and attracting new clients, improving existing solutions for international organizations, and by being able to offer HC-based information and consultancy services accompanying deployed platforms.

Vienna University of Economics and Business (WU)

The *Semantic Knowledge Base* reported in Deliverable D3.3 is a tremendous asset for WU and its research partners. It represents a cornerstone technology that will be used in ongoing and future opinion mining projects, and resonates very well with the *linked data* and *open data* initiatives pursued by WU researchers.

The *Protégé* plugin makes the uComp API available to a wide range of users, and highlights the competence of WU in the areas of ontology engineering and ontology learning. The semantic knowledge base as a central repository to store the acquired knowledge is the logical next step, and will have a strong impact on WU's research roadmap - it will not only play an important role in a planned big data H2020 project proposal to be submitted in in the first half of 2017, but it will also be exploited in the new "CommuniData" open data project, which aims to build an *Urban Participation Platform* to leverage, co-create and discuss *Open Data.*

The platform will support citizens in collective decision making, providing data-rich evidence when engaging in public deliberation processes. uComp's lessons learnt in terms of user engagement will be very valuable when defining the project's incentive strategies, while the semantic knowledge base will be used to host and manage factual (e.g. named entities) and affective knowledge (e.g. sentiment lexicons), and make this knowledge available by means of SPARQL or REST APIs.

Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI)

Information exchange among humans expressed by the means of language is a major research focus of LIMSI. But today's information is becoming more and more "user-generated", and as a result, less structured. The uComp advances in terms of factual and affective knowledge extraction will support analysis and understanding of these various forms of "noisy" texts, while the project's human computation components yield gold standard data and represent an effective mechanism to engage communities around the production and use of language resources. uComp activities related to evaluation and sentiment modelling provided the basis upon which new research activities and collaborations will develop, with as the first step, the planned open shared task on sentiment analysis on German social media scheduled to take place at the end of 2016.

Higher Education is another important task fulfilled by the laboratory, as LIMSI is a host laboratory for several Doctoral Schools from Paris Saclay (SMEMAG, STIC, SSMMH) and from UPMC (SMAER). LIMSI's members play an active role in the doctoral and master degrees, and plan to incorporate the results of uComp into the curricula of the various degree programs.

SWOT Analysis

The following table summarises the results of a SWOT analysis, pointing out internal *strengths* and *weaknesses* of the uComp project, as well as observable external *opportunities* and *threats*.

	Helpful in achieving the objective	Harmful in achieving the objective
Internal origin	**STRENGTHS** Extensive In-house expertise: - ontology learning - human computation - linked and open data - text mining and NLP - knowledge extraction - GWAP development Sufficient hardware resources GATE summer school Experienced project team Effective project management structures	**WEAKNESSES** Inevitable HR fluctuation of project staff (proved to be just a minor issue) - researchers - administration Infrastructure Level - hardware failures Application Level - task complexity - task repetition (mitigated through careful GWAP interaction design)
External origin	**OPPORTUNITIES** Strong international partners: - NOAA CPO - UNEP - WWF, Earth Hour Global GATE open source community Strong demand for human computation: - Collective awareness - Co-creation and innovation - Distributed decision making Strong demand for knowledge extraction: - Big data initiatives - Marketing, Web intelligence - Artificial intelligence research	**THREATS** Infrastructure Level - intrusion attempts, DoS (Denial of Service) Application Level - Cheating attempts - Lack of interest Strong GWAP competition More restrictive policies of social media platforms (mitigated through consistent security policies, game workflow and quality control procedures, and a flexible data acquisition architecture based on REST APIs)

Legal Aspects and IPR Sharing

Special emphasis has been placed on ethical issues of human contributor engagement (T6.2). uComp followed the general code of ethics and conduct of the *British Psychological Society* (BPS),[3] and the *Austrian Data Protection Law.*[4] USFD as a UK project partner is also obliged to comply with the UK Data Protection Act-1998[5] when personal data is stored and processed locally and therefore this data management plan is based these principles but this plan additionally embodies the guiding principles of EU Directive 95/46/EC5.[6]

Online Application Consent

- *Social Media Platforms*. For data that is publicly available, uComp complies with the social media's privacy policy that governs the use of such data. For applications deployed to collect data, we use the social media platform's permission system and offer a Webpage with details about the nature of data being collected, and the purpose of collection.

- *Media Watch on Climate Change.* We provide a signup page that informs the participant about the nature of data collection, and offer regular updates via a newsletter about platform features (which typically do not infringe on any personal rights, since the platform exclusively serves data explicitly marked as being public).

Data Management

User data collected during the project, with the exception of generated language resources that do not contain any personal information, will not be released to any third party and will not be transferred to countries outside the EU. The collected data is stored on secure encrypted drives on servers. Physical access to these servers was restricted to security cleared IT staff. Online access to the data was restricted to individual members of project and uses encrypted SSL communication channels.

Software and Data Resources

Many of the methodological advances of the uComp project have been released in the form of open source libraries, as summarized in the following list. Partners involved in their development will continue to maintain these libraries, and plan to extend them through follow-up research projects, or commercial applications of the developed technologies (via existing IPR arrangements between scientific partners and spin-out companies).

---

[3] www.bps.org.uk/sites/default/files/documents/code_of_human_research_ethics.pdf
[4] www.dsb.gv.at/DocView.axd?CobId=41936
[5] http://www.legislation.gov.uk/ukpga/1998/29/contents
[6] eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML

List of Open Source Libraries

- eWRT - easy Web Retrieval Tookit also at
  http://www.semanticlab.net/index.php/eWRT;

- a part-of-speech tagger for user-generated noisy text
  https://gate.ac.uk/wiki/twitter-postagger.html;

- the TwitIE GATE-based NLP pre-processing application
  http://gate.ac.uk/wiki/twitie.html;

- the GATE Crowdsourcing plugin, including automatic adjudication tools
  https://gate.ac.uk/wiki/crowdsourcing.html;

- A Python implementation of spectral association http://wwwai.wu-
  wien.ac.at/~wohlg/spectral_association;

- a uComp Plugin for Protégé Ontology Editor: The plugin that facilitates the in-
  tegration of typical crowdsourcing tasks into ontology engineering from within
  Protégé;
  https://github.com/UcompWu1/Gwap-Protege-Plugin

- A large, fully annotated named entity recognition dataset for noisy text, con-
  taining ten thousand annotated tweets marked by both crowd and experts, for
  NLP and also crowd analysis;

- A named entity linking dataset for twitter
  http://www.derczynski.com/sheffield/resources/ipm_nel.tar.gz;

- Brown clusters over multiple text types and with various hyperparameter vari-
  ations https://s3-eu-west-
  1.amazonaws.com/downloads.gate.ac.uk/resources/derczynski-chester-
  boegh-brownpaths.tar.bz2;

- Code for generalised Brown clustering, an unsupervised technique for finding
  similar words
  https://github.com/sean-chester/generalised-brown;

- DEFT 2015 Training and Tests Corpora with manual annotation of opinions,
  sentiments and emotions at various granularity levels of Tweets in French
  about climate change
  https://deft.limsi.fr/2015/corpus/train/TRAIN_TWEETS_ID-03042015.zip

- Basic command line Tweet retriever, a simple script to retrieve Tweets by
  their identifiers.
  https://deft.limsi.fr/2015/tools/tweet_basic-retriever.zip;

- DEFT 2015 evaluation toolkit, set of perl programs to compute the DEFT
  2015 evaluation measures
  https://deft.limsi.fr/2015/tools/evaldeft2015_20150513.tar.gz

Indirectly, the generated IPR will also be shared through collaboration with other projects. The *GATE crowdsourcing plugin*, for example, has now been taken up for use, maintenance, and, if needed, further development as part of follow-up research projects. Similarly, the *semantic knowledge base* will be used as a central repository to manage lexical resources and named entities within the InVID Innovation Action.

- SoBigData - an H2020 research infrastructure project on social media analytics, where human computation forms an important part;

- DecarboNet - an FP7 project, which is using the uComp GATE crowdsourcing plugin to create gold-standard datasets for evaluation of entity linking, environmental term extraction, and opinion mining;

- COMRADES - an H2020 CAPS project, which will use the uComp GATE crowdsourcing plugin to create much needed gold-standard datasets for evaluating algorithms for analysing disaster response social media content;

- InVID (In Video Veritas) - a H2020 Innovation Action that will use the semantic knowledge base as a central repository for a verification platform to detect emerging stories and assess the reliability of user-generated video content.

Appendix A: Selected Media Coverage

Greenpeace Magazin, 26/03/2015

"Spielen gegen den Klimawandel",

www.greenpeace-magazin.de/nachrichtenarchiv/spielen-gegen-den-klimawandel

quiz 26.Mär 2015

Spielen gegen den Klimawandel

Crowd vs. Wissenschaft: Mit einem neuen Quiz kann man sein Wissen über den Klimawandel testen und die eigenen Angaben mit den wissenschaftlichen Prognosen abgleichen. Gar nicht so einfach.

Was genau bedeutet Treibhauseffekt? Welchen Anteil hat Kohlendioxid am Klimawandel? Und wie viel Prozent durchschnittliches Schneewasseräquivalent wird die Schneedecke der amerikanischen Sierra Nevada am 1. Mai 2015 enthalten? Wer überzeugt ist, viel über die Erderwärmung und ihre Ursachen zu wissen, wird von dem heute veröffentlichten Klimawandel-Quiz „Climate Challenge" auf die Probe gestellt.

Das Webprojekt entstand in Zusammenarbeit der privaten Modul-Universität in Wien mit der amerikanischen Wetter- und Ozeanographiebehörde NOAA. Während des Spiels muss der Nutzer aus vorgegebenen Antworten die richtige wählen oder auf einer Skala Vorhersagen treffen. Anschließend werden die eigenen Angaben mit den wissenschaftlichen Daten und den Prognosen der Experten abgeglichen. Damit ist Climate Challenge nicht einfach nur ein Quiz, sondern geht der Frage nach, ob das geballte Wissen der Crowd an das der offiziellen Experten heranreicht.

Mit dem Social Media Game „Climate Challenge" testen Nutzer ihr Wissen über den Klimawandel und messen sich so mit Freunden. Foto: Modul-Universität Wien

„Nur zu informieren oder neue Technologien zu entwickeln, reicht nicht als Strategie zur Senkung des weltweiten Energiekonsums", sagt Arno Scharl, Leiter des Instituts für Neue Medientechnologie an der Modul-Universität. „Wir müssen die Öffentlichkeit viel stärker als bisher in Diskussionen zum Thema Klimawandel einbinden".

Auch für die Projektpartner ist das Quiz ein Testlauf. Sie wollen damit untersuchen, wie Social Media zur Sensibilisierung für eine nachhaltige Lebensweise genutzt werden kann. Daher ist das Spiel direkt mit Netzwerken wie Facebook verknüpft – so können sich User mit Freunden und Teilnehmern weltweit messen. Durch eigens entwickelte Visualisierungen wird sofort deutlich, wie sehr die eigene Antwort von der Anderer abweicht. Praktische Tipps zur Reduzierung des eigenen ökologischen Fußabdrucks gibt es als Hintergrundinformationen.

Der Schwierigkeitsgrad der Fragen ist durchmischt, aber bei der Gestaltung der Oberfläche hätten die Macher mehr Liebe investieren können.

Kurt Stukenberg

Krone.at, 27/03/2015

"Wiener Game soll für Klimawandel sensibilisieren",

http://www.krone.at/digital/wiener-game-soll-fuer-klimawandel-sensibilisieren-climate-challenge-story-445588

⌂ Startseite > Digital > Wiener Game soll für Klimawandel sensibilisieren

Climate challenge
test your climate knowledge · predict the future state of our world

Foto: modul.ac.at

"Climate Challenge"

27.03.2015, 06:30

Wiener Game soll für Klimawandel sensibilisieren

Ein Spiel mit Hintergedanken haben Wissenschaftler der Modul University Vienna gemeinsam mit der amerikanischen Klimawandelbehörde NOAA gestartet. Mithilfe des Online-Wettbewerbs "Climate Challenge" sollen User auf spielerisch-kompetitive Weise zum Nachdenken über nachhaltige Lebensweisen angeregt werden.

D abei gilt es, Fragen über zukünftige Klimabedingungen zu beantworten. Eine Visualisierung zeigt, wie sehr die eigene Antwort von der von Freunden, der Community oder Experten abweicht. Die Punkte werden ermittelt, sobald die später tatsächlich gemessenen Umweltdaten verfügbar sind.

Das Lesen von entsprechenden Hintergrundinfos erhöht die Chance, die Antworten zu verbessern. Außerdem winken monatliche Preise und ein Platz in der "Hall of Fame". Tipps zur Reduktion des CO2-Fußabdrucks ergänzen die "Climate Challenge" .

"Wir müssen die Öffentlichkeit viel stärker als bisher in Diskussionen zum Thema Klimawandel einbinden und sie zur Änderung von Verhaltensmustern ermutigen", so Arno Scharl, Leiter des Instituts für Neue Medientechnologie an der Modul University, in einer Aussendung.

Der Anwendung zugrunde liegt das EU-Projekt DecarboNet , in dem erforscht wird, wie soziale Medien zur Sensibilisierung für eine nachhaltigere Lebensweise eingesetzt werden können.

Bluewin.ch, 10/08/2015

"Online Spiel Climate Challenge unterstützt Klimaschutz",

https://www.bluewin.ch/de/leben/nachhaltigkeitsblog/2015/15-08/wir-schweizer-sind-umweltbewusster-geworden.html

Online-Spiel «Climate Challenge» unterstützt Klimaschutz

Mit dem Velo statt dem Auto zu fahren ist auch ein Beitrag zur Bekämpfung des Klimawandels - und macht erst noch Spass (Symbolbild).
Bild: Getty images

Sascha Bianchi
10.08.2015 - 16:45
Bewertung
★★★★★
Bewertungen: 2

Laut einer Studie geht es vorwärts mit dem Umweltbewusstsein der Schweiz. Aber was können wir konkret für Energiewende und Co. tun? Ein von Wissenschaftlern entwickeltes Online-Spiel bietet Unterstützung.

57 Prozent der Schweizer Bevölkerung betrachten sich als überdurchschnittlich umweltbewusst. Das sagt die Studie Univox Umwelt 2014 des Markt- und Sozialforschungsinstitut GFS-Zürich, welche in Zusammenarbeit mit dem WWF Schweiz erstellt wurde. Demnach ist das Umweltbewusstsein in der Schweiz deutlich gewachsen. Fast drei Viertel der 1009 Befragten sind jedoch der Meinung, dass noch deutlich mehr für den Klimaschutz unternommen werden müsste. Für eine Mehrheit (51 Prozent) sollte die Schweiz gar den Anspruch haben, das klimafreundlichste Land Europas zu werden (dagegen sind 23 Prozent).

Energiewende als Vorteil für die Wirtschaft?

Für die meisten Menschen ist die Energiewende dabei nicht nur eine gute Idee, sondern ein umsetzbares Ziel. 15 Prozent erwarten Nachteile für die Wirtschaft, während 44 Prozent die Energiewende als wirtschaftlich vorteilhaft einschätzen. «Für eine Mehrheit sind Umweltschutz und wirtschaftliches Wohlergehen offenbar kein Gegensatz», sagt Andreas Schaub, Geschäftsführer von GFS-Zürich. «Die Energiewende wird gar als wirtschaftliche Chance wahrgenommen.»

«Nur informieren reicht nicht»

Die Energiewende und damit die Bekämpfung des Klimawandels

Bieler Tagblatt, 26/06/2013

"Erfolge im Forschungs-Projekt uComp der MODUL University Vienna",

http://www.bielertagblatt.ch/unterhaltung/digital/erfolge-im-forschungs-projekt-ucomp-der-modul-university-vienna

Bieler Tagblatt

Donnerstag, 14. November 2013

Startseite › Unterhaltung › Digital › Erfolge im Forschungs-Projekt uComp der MODUL University Vienna

SOZIALE MEDIEN   26.06.2013, 11:47

Erfolge im Forschungs-Projekt uComp der MODUL University Vienna

Methoden zur intelligenten und automatischen Gewinnung von Wissen aus sozialen Medien werden derzeit an der MODUL Universität Vienna entwickelt - und wurden jetzt im Vorfeld einer internationalen Konferenz als Open Source Tool online publiziert.

Soziale Medien - Automatisch aber intelligent nutzen

(mt) Die Methoden umfassen die Erfassung und Filterung von Inhalten, Spracherkennung, Fehler-Korrektur sowie eine vereinheitlichte Archivierung der gewonnenen Information. Das Open Source Tool ist ein Beitrag zu einem offenen und transparenten Umgang mit Daten aus sozialen Netzwerken und stellt einen wesentlichen Fortschritt im uComp-Projekt der MODUL Universität Vienna dar. Dieses demonstriert am Beispiel von Klimawandel modernste Methoden zur Analyse von Online-Informationen und kombiniert diese mit kollektiver menschlicher Intelligenz ("Wisdom of the Crowds").

Das Internet ist keine wohlstrukturierte Literatur-Datenbank. Im Gegensatz zu Bibliotheken oder den internen Archiven großer Organisationen liegt Information hier fragmentiert und ungeordnet vor. Das erschwert die automatische Extraktion von Wissen. Noch komplizierter wird das mit dem Siegeszug

Information-age.com,   **09/07/2013**

"Using man and machine to understand the web"

http://www.information-age.com/technology/information-management/123457187/using-man-and-machine-to-understand-the-web--

InformationAge — *Insight and analysis for IT leaders*

Using man and machine to understand the web

A project at MODUL University Vienna aims to combine sentiment analysis, crowdsourcing and gamification to build the ultimate gauge of public opinion

Posted by Chloe Green on 9 July 2013

The worldwide web is an unprecedented source of market intelligence that, in theory, allows businesses to analyse public opinion automatically.

But human communication is complex. Context, slang, varying dialects and other foibles make it extremely difficult for a computer to extract what someone means from what they write online.

Professor Arno Scharl, of MODUL University Vienna's Department of New Technology, has set out to crack this conundrum with an approach that combines automated analysis with human intuition. And, he claims, his team is making significant strides.

Related topics
Analytics

Related articles

European project aims to bring Google-scale web analytics to all

Medianet, 05/07/2013

"Social-Media Daten automatisiert auswerten"

http://medianet.at/front/download_singlepage/5547

Freitag, 5. Juli 2013

FACTS & FIGURES

Modul University Vienna Forscher arbeiten an der automatisierten Extraktion von Wissen aus Sozialen Medien

Social Media-Daten automatisiert auswerten

Wien. Die Methoden zur intelligenten und automatischen Gewinnung von Wissen aus Sozialen Medien, die derzeit an der Modul University Vienna entwickelt werden, sind ab sofort als Open Source Tool online verfügbar. Die Methoden umfassen die Erfassung und Filterung von Inhalten, Spracherkennung, Fehlerkorrektur sowie eine vereinheitlichte Archivierung der gewonnenen Information. Das Open Source-Tool ist ein Beitrag zu einem offenen und transparenten Umgang mit Daten aus Sozialen Netzwerken und gehört zum „uComp-Projekt" der Modul University Vienna. Dieses demonstriert am Beispiel des Klimawandels Methoden zur Analyse von Online-Informationen und kombiniert diese mit kollektiver menschlicher Intelligenz („Wisdom of the Crowds").

Besserer Überblick

„Millionen Menschen äußern ihre Meinung dazu in Sozialen Medien, aber mit herkömmlichen Methoden können wir das gemeinsame Stimmungsbild nicht wahrnehmen",

erklärt Arno Scharl, Leiter des Instituts für Neue Medientechnologie an der Modul University Vienna, das Projekt: „Wir wissen nicht, welche Aspekte berühren, mobilisieren oder zum Nachdenken anregen. Die Technologien von uComp verbessern unsere Möglichkeiten, solche Stimmungen einzufangen – und zwar auf globaler Basis, unabhängig von Sprachbarrieren, Landesgrenzen und Kulturunterschieden."

Der wesentliche Aspekt von uComp ist die Kombination kollektiver menschlicher Intelligenz mit der Fähigkeit zur automatisierten Wissensextraktion durch Software-

© Modul Universität Wien

Arno Scharl, technischer Direktor des uComp-Projekts der Modul University.

Tools. Der erste Schritt dazu ist mit dem sogenannten extensible Web Retrieval Toolkit (eWRT) gelungen. Dieses Open Source Software-Paket wurde nun in einer wissenschaftlichen Publikation vorgestellt und erlaubt es, große Mengen öffentlicher Daten aus Sozialen Medien zu extrahieren und zu analysieren. Unter anderem kann das System die verwendete Sprache präzise identifizieren.

Online-Spiele im Einsatz

In den kommenden zweieinhalb Jahren wird es darum gehen, kollektive menschliche Intelligenz für die erkenntnisorientierte Analyse der mittels eWRT gewonnenen Daten zu nutzen. Eine konkrete Umsetzung solcher als Human Computation (HC) bezeichneten Ansätze sind Online-Spiele zur Klassifikation von digital erfassten Texten oder zur Beurteilung von automatischen Übersetzungen. Mit dem Einsatz zur Identifizierung komplexerer Wissens-Muster betritt das uComp-Projekt digitales Neuland.

www.modul.ac.at/nmt

phys.org, **26/06/2013**

"uComp research project delivers first results under open source license"

http://phys.org/news/2013-06-ucomp-results-source.html

Appendix B: Partner Dissemination

Press Release NOAA, 24/03/2015, "NOAA unveils Climate Challenge: an interactive game to build and test climate knowledge", cpo.noaa.gov/OutreachandEducation/OutreachArchive/TabId/505/ArtMID/1254/ArticleID/242371/NOAA-unveils-Climate-Challenge-an-interactive-game-to-build-and-test-climate-knowledge.aspx

NOAA CLIMATE PROGRAM OFFICE
Advancing scientific understanding of climate, improving society's ability to plan and respond

Home   About CPO   Climate Programs   Grants and Projects   Outreach and Education   Partnerships

Outreach and Education   Outreach Archive

NOAA unveils Climate Challenge: an interactive game to build and test climate knowledge

Have you ever wanted to test your climate prediction skills against the experts'? Now you can! Today, the National Oceanic and Atmospheric Administration (NOAA) Climate.gov team released *Climate Challenge,* a new educational online game that invites players to learn more about climate science and to test their knowledge. The game is available at: www.ecoresearch.net/climate-challenge.

Climate challenge
test your climate knowledge · predict the future state of our world

Climate Challenge allows players to guess future climate-related conditions, such as the annual Arctic Sea Ice minimum coverage or a monthly average global surface temperature. The game also allows players to compare their answers to the experts' and, when available, see how both predictions compare to real-world measurements. There will be one climate-related question per month, and whoever comes closest to guessing the actual values will be that month's winner. For an added twist, participants can also compete against their friends, colleagues, or family members.

Climate challenge
test your climate knowledge · predict the future state of our world

Play   Progress   Leaderboard   About

El Niño-Southern Oscillation (ENSO)   Deadline: March 30, 2015

How much warmer or colder than average will sea surface temperature be in April 2015 in the eastern tropical Pacific (the Niño 3.4 region)?

0 °C (0 °F)   Submit
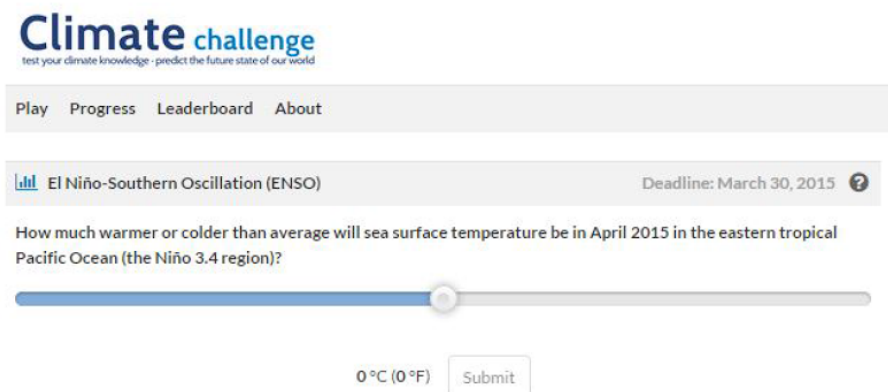
The game was inspired, in part, by the *New Yorker* business columnist James Surowiecki's best-selling book titled *The Wisdom of Crowds*, which explores the concept that large groups of people all acting as individuals are, collectively, smarter and better at solving problems, fostering innovation, coming to wise decisions, and even predicting the future than an elite few — no matter how brilliant the elites may be.

MOD Website

WU Website

Press Release MU Vienna, 26/06/2013, "Opinion Lead versus Information Overload:
Automatic – But Smart – Use of Social Media",

www.modul.ac.at/article/view/opinion-lead-versus-information-overload-automatic-
but-smart-use-of-social-media/

PRESS RELEASE

Opinion Lead versus Information Overload: Automatic – But Smart – Use of Social Media

MODUL University Vienna's uComp research project delivers first results under an open source license

Vienna, Austria 26 June 2013 – Methods to extract knowledge from social media intelligently and automatically are currently being developed at MODUL University Vienna – and the latest advances have just been published in preparation of an international conference. These advances come in the form of an open source tool to collect and process publicly available social media information. The tool supports text acquisition, language recognition, detection of phonetic similarities, as well as the standardized integration and archiving of the captured information. The open source tool represents a major step forward in the uComp project of MODUL University Vienna (Austria) and its European partners. Using the domain of climate change as an example, the project combines cutting-edge methods to automatically capture information from complex sources and combine it with collective human intelligence in the tradition of the "wisdom of the crowds".

The Internet is very different from a well-structured database. Unlike libraries or large corporate archives, online information is fragmented and unordered, which makes it difficult to extract knowledge automatically. The emergence of social media has further complicated the process. It is difficult to determine the specific context of a posting, and the use of slang, dialects or foreign words challenges existing tools for text analysis. Scientists and researchers are currently working on solving this problem in the uComp project jointly conducted by MODUL University Vienna and partner organizations from Austria, England and France. After only six months, first results have now been published in preparation of the *7th International Conference for Knowledge Capture* (K-Cap 2013) in Banff, Canada.

Man/Machine Symbiosis

The objective of uComp is explained by the head of the Department of New Media Technology at MODUL University Vienna, Prof. Arno Scharl, using the domain of climate change as a use case: "Millions of people express their opinions in social media, but with conventional methods we are unable to determine the collective mood expressed in social media in real time. We do not know which aspects move people, mobilize people or stimulate their thoughts. The technologies from the uComp project provide us with better ways to capture opinions – on a global basis, irrespective of language barriers, national borders and cultural differences."

The key aspect of uComp for Prof. Scharl, who also serves as the project's Technical Director, is the combination of collective human intelligence and automated knowledge extraction by software tools. The first step to achieving this vision has successfully been taken with the