Social Media Data analyzis and Semantics for Tourism Understanding

Jérôme Treboux, Fabian Cretton, Dominique Genoud , Anne Le Calvé,
Florian Evéquoz, Antonio J. Jara[1], Blaise Larpin[2]

[1] Institute of Information Systems, [2] Institute of Tourism, HES-SO Valais, Switzerland

ABSTRACT

How social media could be used to interpret the satisfaction of clients visiting a destination based on real use cases?

- **2.078 Billion** active social media accounts (**29%** of the total population)
- **+12%** active social media account compared to 2014 **(+222** Million)
- **16 min/hour** is the time spent on social networks in the USA
- **23%** of Facebook users login at least **5** times a day [1, 2]

More than just to communicate with clients, this analysis let the resort analyzing the effective tourist needs and hope when he is coming in this tourism hotspot (Mountain Bike, Ski, …).
With a semantic approach, it is possible to know what the interests of tourists are when they are traveling in a specific region.

Fig. 1 – Work flow from the social network to the client

RESEARCH DESIGN AND METHODOLOGY

1. Use cases
   - Partnership with a company active in the tourism
   - Orientation of the project
   - *"Based on the Social Media Data and pictures taken by tourists, the Tourism offices should be able to Identify the Tourist's Needs and Adapt their marketing to save time and money"*
2. First analyzis
   - Focus on Instagram
   - Retrieving all pictures from Instagram through a KNIME workflow
   - Confirmation of the tourists presence over social media

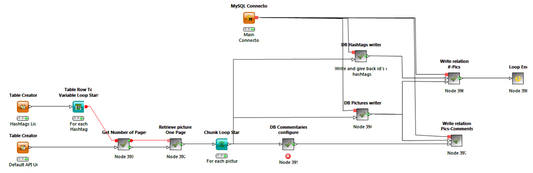

Fig. 2 – KNIME Workflow retrieving pictures from Instagram

3. Data preparation
   - Tags extraction from the tag field and comments over the picture
   - Cleaning Process of the data
4. Semantic
   - English, French, German
   - BabelNet Ontologies usage [3, 4, 5]
   - 13'000 tags linked to a concept

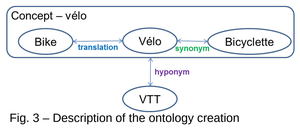

Fig. 3 – Description of the ontology creation

5. Visualization
   - Result of the project for the client

SYSTEM MODELING AND RESULTS

Visualization: "the use of computer-supported, interactive, visual representations of abstract data to amplify cognition" [6]

Objectives:
- Identify the Tourist Needs
- Adapt the Marketing

1. Visualization of Origin and Destination

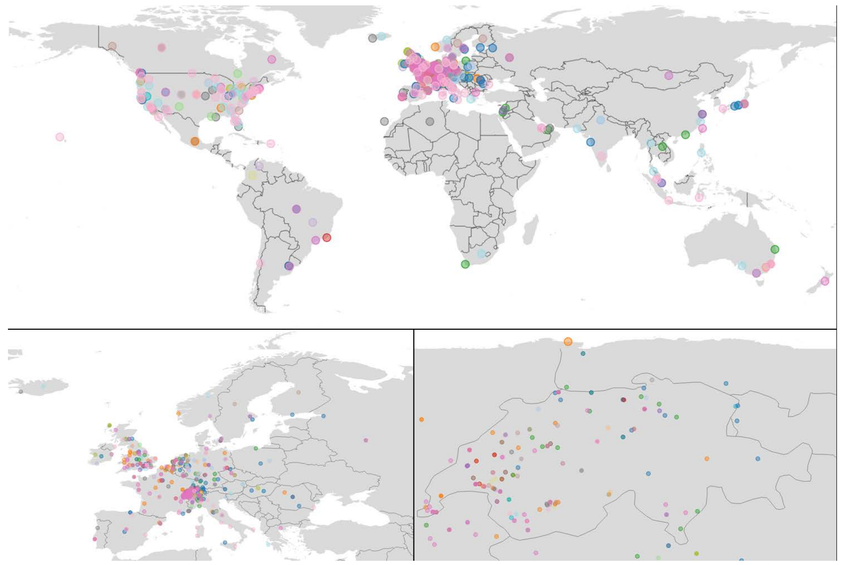

Fig. 4 - Origin of tourists visiting Valais according to Flickr. Position of dots correspond to origin of tourist. Color corresponds to a destination. World map (top), European map (bottom left), Swiss map (bottom right).

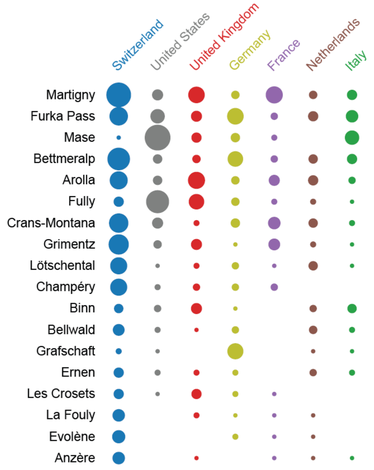

Fig. 5 - Tabular visualization of pictures from Valais posted on Flickr. Each bubble represents the pictures taken at a specific destination (row) by people coming from a specific country of origin (column and color). The area of the bubble is proportional to the number of pictures.

2. Visualization of Semantic Data

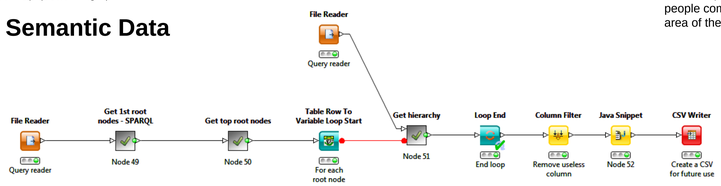

Fig. 6 – KNIME Workflow – Visualization for graph the graph exploration tool

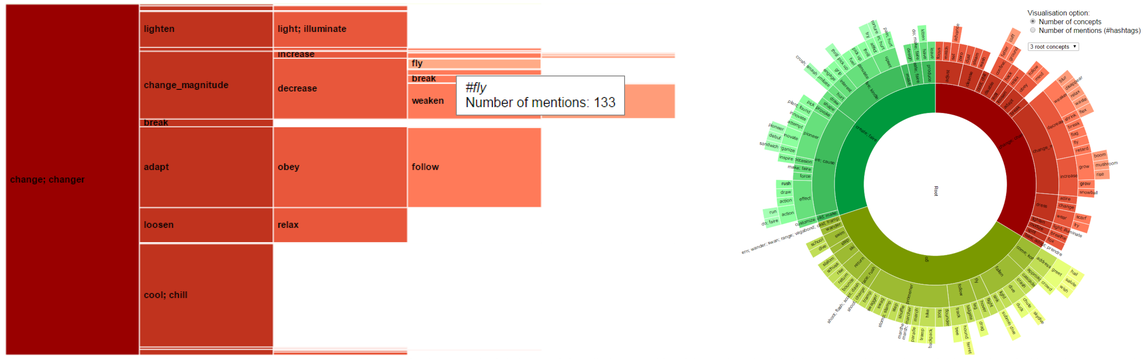

Fig. 7 – Graph Exploration Tool result. The tool provides two zoomable visualizations. On the left the graph is called "Zoomable Icicle" and on the left, it is a "Zoomable sunburst" both developed with D3.js, a JavaScript library for manipulating documents based on data

CONCLUSIONS

The goal to find what Tourists are thinking and looking for in a resort is reached

Benefits:

1. The tourism office is able to adapt their communication through the website or directly on social networks
2. This tool provides this analyzis on real time data
3. Time and money saving by avoiding traditional survey and use social media to refine the customer knowledge

FUTURE WORK

Some optimization of the tool to be completely usable in a production mode:

1. Adding different social media sources
2. Searching on social media though a semantic process to be more accurate on results
3. Location of the picture (GPS, image recognition, …) to check that it is from the right place
4. Disambiguation of words retrieved
5. Improvement of the visualization based on client feedback

REFERENCES

1. Kemp, S., Digital, Social & Mobile Worldwide in 2015,We Are Social, http://wearesocial.net/blog/2015/01/digital-social-mobile-worldwide-2015/, January 2015
2. Casas, J., Social Media Trends 2014: The latest data, infographics and statistics, postcron.com, http://postcron.com/en/blog/social-media-trends-2014/, 2015
3. Berners-Lee, T., Linked data-design issues, 2006
4. Bizer, C., Heath, T. & Berners-Lee, T., Linked data-the story so far, 2009
5. Navigli, R. & Ponzetto, S. P., BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network, Artificial Intelligence, 193, 217-250, 2012
6. Card, Stuart K. and Mackinlay, Jock D. & Shneiderman, Ben, Readings in Information Visualization: Using Vision to Think, 1-55860-533-9, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999

ACKNOWLEDGEMENT

This Project is supported by the Institute of Informations Systems and the Institute of Tourism, University of Applied Sciences and Arts Western Switzerland and the Valais Wallis Promotion
Also Ontotext and Roberto Navigli and colleagues (Sapienza University of Rome) for providing a research license for OWLIM and for BabelNet.