Effective and Efficient Online Communication
The Channel Model

Anna Fensel, Dieter Fensel, Birgit Leiter and Andreas Thalhammer
Semantic Technology Institute (STI) Innsbruck, University of Innsbruck, Technikerstraße 21a, A-6020 Innsbruck, Austria

Keywords: Social Media, Web 2.0, Semantic Web, Dissemination, Communication, Knowledge Management.

Abstract: We discuss the challenge of scalable dissemination approach in a world where the number of communication channels and interaction possibilities is growing exponentially, particularly on the Web, Web 2.0, and semantic channels. Our goal is to enable smaller organizations to fully exploit this potential. We have developed a new methodology based on distinguishing and explicitly interweaving content and communication as a central means for achieving content reusability and thereby scalability over various, heterogeneous channels. Here, we present in detail the communication channel model of our approach.

1 INTRODUCTION

Fax, phone, and later the Internet, have radically changed our communication possibilities. More and more communication has been freed from the geographical barriers that formerly limited their speed and expansion. Now, it is (in principle) possible to instantly communicate with a large portion of the entire human population. Nevertheless, new means also generate new challenges.

Take the world of the TV consumer as an example. Twenty-five years ago, there were around three channels. Therefore, selecting your program was a rather trivial task which required no more than a few seconds. Whilst hundreds of channels have been added, thousands of channels have been connected via the Internet, where extremely large libraries of videos (which go beyond the metaphor of a 'channel'), currently define the content. Now the consumer could spend the rest of his/her life browsing in search of the program he/she wishes to watch. Obviously, he/she requires either new skills or new access means to scale and filter the exponentially increased information service offer. Or assume the task of a small hotelier. How can it be ensured that the hotel is found by potential customers, i.e., how can he/she find them? It should have a web site with high visibility by various search engines, and must be present in a large number of on-line booking channels, we should find it through the town's website, and obviously a Facebook site is a must (with a booking engine included). Bookings through mobile platforms are significantly increasing and the hotelier would want to be found there too. Why not add a video about the hotel on YouTube, a chat channel for instant communication, fast email and fax response capabilities, the old-fashioned telephone, and occasional tweets and emails that are clearly distinguishable from spam? Preferably, the communication should be multi-directional, i.e., the organisation should realize when one of its posts gets commented on (up to a full-fledged impact analysis), or even more importantly, it should be aware when someone talks about the hotel and how much the costumer liked it. As much as this is needed, this obviously does not scale, especially for SMEs with limited marketing activities budgets, and (Mulpuru et al., 2011) calls it "*the growth of the multichannel monster*".

Therefore, organizations of all sizes, commercial and not-for-profit, regularly face the challenge of communicating with their stakeholders using a multiplicity of channels, e.g. websites, videos, PR activities, events, email, forums, online presentations, social media, mobile application, and now, structured data. The social media revolution has made this job much more complicated, because:

- the *number of channels* has grown exponentially,
- the communication changes from a mostly unilateral "push" mode (one speaker, many listeners) to an increasingly fully *bilateral communication*,

Fensel A., Fensel D., Leiter B. and Thalhammer A..
Effective and Efficient Online Communication - The Channel Model.
DOI: 10.5220/0004011402090215
In *Proceedings of the International Conference on Data Technologies and Applications* (DATA-2012), pages 209-215
ISBN: 978-989-8565-18-1

where individual stakeholders (e.g. customers) expect one-to-one communication with the organization, and the expected speed of reaction shrinks to almost real-time, and

- the *contents of communication becomes more and more granular* and increasingly dependent on the identity of the receiver and the context of the communication.

Organizations need an integrated solution that provides the management and execution of communication goals in a mostly automated fashion, at the costs of mass-media communication, with the granularity of individual experts, at the pace of real-time social media. We are aiming to mechanize important aspects of these tasks, allowing scalable, cost-sensitive, and effective dissemination for small-or-medium sized business units and comparable organizations, for which dissemination is essential but resources are significantly limited. Additionally, it may also help intermediaries such as marketing agencies to extend their business scope by increasing the cost-effective ratio.

The core idea of our approach is to introduce a layer on top of the various Internet based communication channels that is domain specific and not channel specific. So one has:

- **information models** (specific for certain domains), that define the information or knowledge items;
- a **channel model** (the communication model), that describes the various channels and their target groups; and
- **mappings** of information items on channels through weavers.

In this paper we present in detail the channel model of our approach.

The paper is structured as follows. Section 2 discusses the channel model of the approach. Section 3 discusses related work, Section 4 sketches future work and Section 5 provides the conclusions.

2 CHANNEL MODEL

According to Wikipedia, "in telecommunications and computer networking, a communication channel, or channel, refers either to a physical transmission medium such as a wire or to a logical connection over a multiplexed medium such as a radio channel." In online communication, we take a broad definition of a channel. A channel is a means of exchanging information in the on-line space. There is a close relationship between URIs and channels as each URI can be used as a channel to spread or access information. However, not each channel directly refers to an URI. For example, Facebook provides around forty different methods of spreading information not distinguished by a URI. Additionally, individual information items spread through Facebook are not distinguished by URIs. In general, a channel can be interpreted as a "place" where one can find or leave information, whether it is unanimously referred by a URI or addressed through a service. However, even this is not broad enough. As described previously, a channel can also be the URI of a vocabulary (or the formalisms such as RDFa or microformats) that are used to publish the information. Through use of this URI, only humans or software agents that "speak" this dialect are able to access this information. Here, the communication channel cannot be interpreted as a place, but rather as a way to express or refer to the information. In the following, we want to distinguish channels by the communication mode they support.

Communication is based on the broadcasting of information. Therefore, we define the first category of our channel classification system as channels used for *broadcasting*. Here we make a distinction between the publication of mostly *static information* and *dynamic contents* that express the timeliness of an information item. One way of spreading information is to invite other people to use it. Therefore, *sharing* is another category we have identified. It reflects the insight that others are not passive consumers of our information but active prosumers that should be helped and supported in their information processing activities. Sharing is the first form of cooperation. Explicit *collaboration* through a shared information space is the next cooperation category we have identified. Collaboration between individuals leads to groups of people actively organizing their communication and cooperation. Social networking sites that support *groups of people* in their information needs are instances of this next category we have identified. Obviously, the boundaries between these categories are fluid and many channel providers try intentionally to establish services covering several of them. Still, it is often possible to identify a major category for them, often based on the major usage patterns of their users. An important approach to broaden the scope of a dissemination activity is to add machine-processable semantics to the information. With this approach, search and aggregation engines can provide a much better service in finding and retrieving this information. A

means of adding *machine-processable semantics* to information is our final channel category.

Broadcasting Static Information. Websites are an established means of providing (mostly) static information. Information that reflects the structure of the contents is provided through websites and they offer a smooth way for users to access this content. An important addition beyond the dissemination through an owned website is an entry on other sites such as Wikipedia, the world's leading encyclopedia.

Broadcasting Dynamic Information. With Web 2.0 technologies, dedicated means for publishing streams and interacting with information prosumers have been added. A first step in this direction is the inclusion of a News section in a website using blogging tools such as Wordpress. Good practices for a news section on a website are:

- Each news item has its own URL, so that they can be returned in search results, bookmarked, shared etc.;
- News should contain a pointer to a more detailed description about the information items they describe;
- each news item is archived;
- each news item can be indexed by search engines;
- each news item is typed (through use of the information model);
- each news item is categorized (through use of a folksonomy);
- each post can be directly shared, emailed, added to favorites, and liked;
- news can be searched, sorted, and filtered; and
- important news items stay at the top to highlight main announcements.

Such news can be further spread through a news ticker such as *RSS feeds* and *Twitter***.** An RSS feed is used to broadcast news. Its purpose is to regularly remind the user of the existence of a particular activity and the fact that it is producing interesting results. Twitter is a widely used means of disseminating news, however, significantly limits the length of it. Finally, *Email* and *Email lists* are also well established means for news dissemination. Especially the latter are a proven means of broadcasting information and facilitating group discussions. Other ways of spreading news are through social networking sites, which will be discussed below. *Chatting* is another form of instantly communicating and disseminating information, and a *blog* could be used to inform partners and members of recent trends in the field of semantic technologies.

Sharing. There are a large number of Web 2.0 websites that support the sharing of information items such as: bookmarks, images, slides, and videos, etc.

Collaboration. A *wiki* is primarily a means for project internal collaboration. However, it also becomes a dissemination channel if external visitors have *read* access. They may then follow the intensive internal interaction that can help to gain a better and more detailed understanding of externally published results and achievements.

Group Communication. *Facebook* as a social networking site provides an additional community aspect, i.e., it forms a community that multi-directionally shares news, photos, opinions, and other important aspects. Notice that Facebook is actually not only one, but several channels. It offers more than 40 possibilities through which to disseminate information. These can also be tightly integrated into Web 1.0 pages, such as that of the New York Times. *Google+* may have the potential to become a major competitor of Facebook. Therefore, it should also be included in a social networking site strategy. *LinkedIn* and *Xing* are focused on professional use and perfectly fit the purpose of research organizations.

Semantic-based Dissemination. An important approach to broaden the scope of a dissemination activity is to add machine-processable semantics to the information. With this approach, search and aggregation engines can provide much better service in finding and retrieving this information. Semantic annotations injected in websites are used by search engines such as Google to provide a structured presentation of the contents of websites. "This data may be embedded within enhanced search engine results, exposed to users through browser extensions, aggregated across websites or used by scripts running within those HTML pages." (Tennison, 2012).

There are various *formats* of adding machine-processable semantics to data. First, there are three competing means of including semantics directly in HTML/XML files: (1) RDFa adds a set of attribute-level extensions to XHTML enabling the embedding of RDF triples; (2) Microformats directly use meta tags of XHTML to embed semantic information in web documents; (3) Microdata use HTML5 elements to include semantic descriptions into web documents aiming to replace RDFa and

Microformats. For the moment, we have three competing proposals that should be supported in parallel until one of them can take a dominant role on the web. RDFa integrates best with the W3C metadata stack built on top of RDF. However, this also seems to hamper the uptake of this technology by many webmasters that are not familiar with this technology stack. Therefore, Microformats were developed as a competing approach directly using some existing HTML tags to include meta data in HTML documents. Actually, they overload the class tag which causes problems for some parsers as it makes semantic information and styling markup hard to differentiate. Therefore, Microdata instead introduce new tag attributes to include semantic data into HTML. As the use of RDFa has increased rapidly, whereas the deployment of microformats in the same period has not advanced remarkably, we are focusing on RDFa and Microdata in our dissemination approach.

Instead of including semantic annotations in XHTML documents, i.e., injecting machine-readable contents into content that is meant for direct human consumption, they can also be provided for direct machine consumption. A straight-forward way is to publish an RDF file containing the machine readable data. Instead of directly publishing an RDF file you can also provide a SPARQL endpoint allowing the querying RDF information. Instead of retrieving the entire RDF file, directed queries can be supported with this approach

In addition to predefined formats and technical means, we need to reuse predefined *LOD vocabularies* to describe our data to enable semantic-based retrieval of information. Currently, we use Dublin Core, FOAF, GoodRelations, and schema.org.

3 RELATED WORK

Here we relate and compare to the two closest sub-areas of semantic technologies: *ontology-based content management systems (CMSs) for web sites* and s*emantic matchmaking of content sender and receiver*.

The field of *semantics-based or enhanced CMSs* has already been quite thoroughly explored. One of the earlier approaches to ontology-based website management is the OntoWebber system described in (Jin et al., 2001). The proposed three-way approach of “explicit modeling of different aspects of websites", “the use of ontologies as foundation for web portal design", and “semi-structured data technology for data integration and website modeling" presents an early but comprehensive approach to semantifying CMSs. OntoWebber introduces an integration layer which adapts to different data sources. In contrast, our approach adapts to different channels rather than to different information sources. A year later, Sheth et al. (Sheth et al., 2002) introduced the SCORE system which defines four key features: semantic organization and use of metadata, semantic normalization, semantic search, and semantic association. Although written in the early days of the Semantic Web, the paper covers topics such as metadata extraction from unstructured text and automatic classification that may also become relevant to our approach. Continuing their earlier work on the Knowledge Web portal, (Garcia et al. 2008) introduce “The Rhizomer Semantic Content Management System” which integrates services with metadata browsing, editing, and uploading. (Corlosquet et al, 2009) proposes a Linked Data extension for Drupal that enables content annotation with RDFa and provides a SPARQL endpoint. The British national broadcaster BBC started to integrate semantic technologies (i.e. Linked Data) in 2009 in order to integrate various data and content sources distributed throughout the enterprise (Kobilarov et al., 2009). As a result, as reported in (Bishop et al., 2010), BBC's World Cup 2010 site is based on semantic repositories that enable the publishing of metadata about content rather than publishing the content itself. While the data input is fixed, different schemas for the output are defined. However, as only one channel for output is considered, the mapping performed is quite straightforward. In contrast, our system accounts for different information needs of various and heterogeneous channels and therefore enables the distribution of content through different portals.

Semi-automatic matchmaking is a well studied field in Artificial Intelligence and related areas. Obviously, we can only select a small sample of approaches in this area, which focus on matchmaking in regard to content. (Katzagiannaki & Plexousakis, 2003) presents a selective information dissemination system that is based on semantic relations. In their paper, the terms in user profiles and terms in documents are matched through semantic relations that are defined using a thesaurus. Similarly, the approach taken by (Morales-del-Castillo et al., 2009) introduces selective dissemination of information for digital libraries based on matching information items to user profiles. Obviously, user profiles correspond to our

channels, however, we instead manually model their relationship with contents. The system introduced in (Ma et al., 2006) uses RDF, OWL, and RSS in order to introduce an efficient publish/subscribe mechanism that includes an event matching algorithm based on graph matching. Our approach, in contrast, matches information items to channels rather than events to users. Also, instead of graph matching, we use predefined weavers for channel selection. While (Morales-del-Castillo et al., 2009) uses fuzzy linguistic modeling and NLP techniques for semiautomatic thesaurus generation and perform a matching based on statistical analysis, we use semantics to manually define the connections between information items and the channels.

Since we aim for high precision and professionalism in on-line communication, we see little use for statistical based semantic methods (natural language understanding, information extraction, etc.). We want to allow the user to abstract from the content-channel level to the content level. However, as we expand towards a full-fledged value management approach that monitors the entire web space for important statements, such methods will be needed. Fortunately, a large number of such web analytical toolkits already exist (Kasper et al., 2010).

4 FUTURE WORK

Introducing a semantic layer on top of communication channels is required to enable a framework that allows common value management. This combination of research fields opens a broad variety of new challenges yet to be solved, in particular, as follows:

Modeling and Interweaving Feedback. Feedback is an important part of all effective communication. Without feedback, the sender - the one who intends to convey information - has no means to validate whether or not the recipient received or understood the message. It is also often preferable to have a fully-fledged two-way conversation instead of simple one-way broadcasting. The Web 2.0 revolution made it ridiculously easy for everybody to use the Internet as a two-way conversation platform where they can provide feedback as well as react to what was said. Therefore, it will be necessary to model feedback and interweave it with content items that we previously published.

Modeling Target Groups. Companies that pursue common value management usually have a very restricted target group of people they want to address. So far in our channel model we do not distinguish between different target groups in different channels. However, different target groups reside on different communication platforms, even though there is some overlap. For example, you will find more young and hip people on Facebook, and more professional users on Xing or LinkedIn, but there are quite a few users that have a profile on both platforms. Nonetheless, they expect a different way of being engaged in different platforms.

Adapting Content. Adapting content is a two part problem (both of which can be commonly solved): first, converting an information item into different formats or, second, automatable transformations such as extracting images, videos or extracting and shortening web links from piece of content, and transforming multimedia content into a different format. However, adapting content in a way that requires creativity and human intelligence is still a strenuous problem that reaches the borders of computability. Examples of such adaptations are shortening or translating an essay, or rewriting a text in a way that matches the target group it addresses.

Crowdturfing and Trust. Crowd sourcing solutions are very capable of dealing with hard to compute problems. However, this also allows for harming platforms that only defend themselves against automated attacks. Malicious activities, such as shaping opinions of a large number of people via social media platforms, and the use of crowd sourcing platforms are becoming more and more popular (Wang et al., 2011). That is why some mechanisms of trust have to be introduced, which is also tightly connected to reputation management. A non-trustworthy source may communicate anything they want, the effect will be very little and its value drastically decreased.

Quality Management. An important part of targeted communication is assessing and improving the quality of conveyed content. Whereas trust, reputation and brand management are influenced by how information is perceived, quality assurance is an inbound process. The business processes for quality management and what this actually mean have yet to be defined for common value management. The bigger the campaign is, the more visible the effect of proper quality management.

Quantification of Brand and Reputation. Similar to the quantification of social values, brand and reputation also have to get a countable unit. It will be very challenging to find fitting metrics, since measures already in existence such as brand equity

are considered meaningful by a small share of marketing professionals (Farris et al., 2010). The combination with social media and the possibilities of sentiment analysis allows more suitable metrics to be introduced next to the existing ones.

Enrichment of Yield Management. Yield management is based on statistical analysis of different parameters, such as pricing, capacity, and demand (Kimes, 2000). Already established calculation models can be extended by channel based reputation, brand value and other, yet to be introduced criteria. For example, in a communication channel where a product's brand (or the product itself) has little reputation and is badly represented, the price of the delivered product could be less than in other channels where it is reputable.

5 CONCLUSIONS

As a main contribution of this paper, we have developed a channel model for an approach to efficient and effective communication, applicable to the multi-channel dissemination conditions of Web 2.0 and Web 3.0 a.k.a. Semantic Web. The approach as a whole facilitates and automates communication for the communication managers, saving them time and money, as well as reducing the information overload for the information consumers.

The following core features characterize our channel modeling principles:

- We use ontologies to model content in order to have a representation layer independent from the communication channel. We want to achieve reuse of content over channels, allowing small organizations to deal with an increasing number of communication channels and exploit their potential. These ontologies model certain vertical domains such as research projects, associations, accommodations, restaurants, bars, touristic events and services, etc. Thus, the ontologies and their channel alignments can be reused at large scale, providing a quick return of the investment necessary to build and maintain them.
- Our approach is bi-directional, i.e., in the same way that we disseminate through concepts we use these concept to aggregate feedback and impact found in various channels.
- We support in an integrated fashion the dissemination via traditional web channels, web 2.0 means, and semantic based channels, using various formats and vocabularies.

In order to broaden the proof of our claims on scalability and reusability of content and ontologies, we are currently performing case studies. First, we use our approach in the dissemination of other research projects and associations. Second, we are entering more commercial areas such as eTourism, where the demand for a scalable dissemination strategy is supported by millions of hotels, given the fact that soon the majority of room bookings will be done on-line.

ACKNOWLEDGEMENTS

This work has been partially supported by EU projects RENDER and PlanetData.

REFERENCES

D. C. Barlund: A transactional mode of communication. In C. D. Mortensen (eds.), *Communication Theory* (2nd ed.), New Jersey Transaction, 2008.

B. Bishop, A. Kiryakov, D. Ognyanoff, I. Peikov, Z. Tashev, and R. Velkov: OWLIM: A family of scalable semantic repositories, Technical report, 2010.

S. Corlosquet, R. Delbru, T. Clark, A. Polleres, and S. Decker: Produce and consume linked data with Drupal, *Springer Constraints Journal*, 1380:763-778, 2009.

P. W. Farris, N. T. Bendle, P. E. Pfeifer, and D. J. Reibstein: Marketing Metrics: The Definitive Guide to Measuring Marketing Performance (2nd Ed.), Pearson Education, 2010.

R. Garcia, J. M. Gimeno, F. Perdrix, R. Gil, and M. Oliva: The Rhizomer Semantic Content Management System. In *Proceedings of the 1st World Summit on The Knowledge Society*: Emerging Technologies and Information Systems for the Knowledge Society, *WSKS'08*, pages 385-394, Springer, 2008.

Y. Jin, S. Decker, and G. Wiederhold. Ontowebber: Modeldriven ontology-based web site management. *In Proceedings of the Semantic Web Working Symposium (SWWS)*, Stanford University, 2001.

H. Kasper, M. Dausinger, H. Kett, and T. Renner: Marktstudie Social Media Monitoring Tools, *Frauenhofer IAO Studie*, 2010.

I.-E. Katzagiannaki and D. Plexousakis: Information dissemination based on semantic relations. In *CAiSE Short Paper Proceedings'03*, 2003.

S. E: Kimes: A strategic approach to yield management. In A. Ingold, U. McMahon-Battie, and I. Yeoman (eds.), *Yield Management: Strategies for the Service Industry*, International Thomson Business Press, 2000.

G. Kobilarov, T. Scott, Y. Raimond, S. Oliver, C. Sizemore, M. Smethurst, C. Bizer, and R. Lee: Media meets semantic web: how the BBC uses dbpedia and linked data to make connections. In *Proceedings of*

European Semantic Web Conference (ESWC 2009), LNCS 5554, pp. 723-737, 2009.

J. Ma, G. Xu, J. L. Wang, and T. Huang: A semantic publish/subscribe system for selective dissemination of the RSS documents. In *Proceedings of the Fifth International Conference on Grid and Cooperative Computing (GCC'06)*, pp. 432-439, 2006.

J. M. Morales-del-Castillo, R. Pedraza-Jimenez, A. A. Ruiz, E. Peis, and E. Herrera-Viedma: A semantic Model of Selective Dissemination of Information for Digital Libraries, *Information Technology and Libraries*, 28(1):21-31, 2009.

S. Mulpuru, H. H. Harteveldt, and D. Roberge: Five Retail eCommerce Trends to Watch In 2011, *Forrester Research Report*, January 31, 2011.

A. Sheth, C. Bertram, D. Avant, B. Hammond, K. Kochut, and Y. Warke: Managing semantic content for the web, *IEEE Internet Computing*, 6:80-87, July 2002.

J. Tennison: HTML Data Guide, W3C Editor's Draft 02 March 2012.