STI INNSBRUCK

ON-LINE VALUE MANAGEMENT

Dieter Fensel, Birgit Leiter, Stefan Thaler, and Andreas Thalhammer

STI Innsbruck, University of Innsbruck,
Technikerstraße 21a, 6020 Innsbruck, Austria
firstname.lastname@sti2.at

SEMANTIC TECHNOLOGY INSTITUTE INNSBRUCK

STI INNSBRUCK
Technikerstraße 21a
A - 6020 Innsbruck
Austria
http://www.sti-innsbruck.a

Tuesday, April 03, 2012

On-line Value Management

D. Fensel, B. Leiter, S. Thaler, and A. Thalhammer

STI Innsbruck, University of Innsbruck, Austria

Abstract. We discuss the challenge of scalable dissemination and communication approaches in a world where the number of channels is growing exponentially. The web, Web 2.0, and semantic channels have provided a multitude of interaction possibilities providing significant potential for yield, brand, and general reputation management. Our goal is to enable smaller organizations to fully exploit this potential. To achieve this, 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.

1 Introduction and Motivation

The last two hundred years have revolutionized international transport and communication. Fax, phone, and later the Internet, have radically changed our communication possibilities. More and more communication has been freed from the geographic 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. The consumer could now spend a lifetime in search of a program he or she wishes to watch. Obviously, consumers require new skills and more efficient access means to scale and filter the exponentially increased offer.

Precisely the same is needed for our overall approach to on-line (or Internet-based) communication. 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? The hotelier should have a website with high visibility on various search engines and must be present in a large number of on-line booking channels. We should find the hotel on the town's website, and the hotel should have a Facebook page, perhaps with a booking engine included. Bookings made through mobile platforms are increasingly popular, 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 hotelier should realize when one of his posts gets commented on (up to a full-fledged impact analysis), or even more importantly, the hotelier should know when someone talks about the hotel, and how much the costumer liked it. As much as this is needed, this obviously does not scale and [Mulpuru et al., 2011] calls it "*the growth of the multichannel monster*". In principle, the hotelier is presented with three equally problematic alternatives:

- He does it on an ad-hoc basis by himself, as a side activity. This would work, however, the number of potential customers (and therefore business opportunities) that he is missing may be tremendous and could, in the long run, take him out of the market completely.
- He builds up a professional communication team by hiring a large number of social media experts and assigns them to manage his various communication channels. Actually, a large hotel chain may be able to do this. In the case of our hotelier, he would find himself even more quickly out of the market due to the high costs attached to this "solution".
- Finally, he could start to cooperate with an external marketing agency. This marketing agency must understand the domain (tourism, accommodations) and the various communications means available to disseminate the contents about our hotel in an effective and efficient fashion. These agencies have some IT support that supports multi-channel dissemination, however, they have to manually adopt, align, and define the content for these channels. In summation, these services are costly and only partial solutions (to limit the high costs of manual labor by dissemination experts).

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 applications, and recently structured data. The social media revolution has made this job much more complicated, because:

- the *number of channels* has grown exponentially,
- the communication has changed from a mostly unilateral "push" mode (one speaker, many listeners) to an increasingly fully *bilateral communication*, where individual stakeholders (e.g. customers) expect one-to-one communication with the organization, and the expected speed of reaction is shrunk 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 management and execution of communication goals in a mostly automated fashion, with costs equivalent to mass-media communication, along with the granularity of individual experts, and at the pace of real-time social media. We are aiming to mechanize important aspects of these tasks, allowing scalable, cost-sensitive, and effective communication for small-

or-medium sized business units and comparable organizations for which information 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.

Communication is a means to an end. Section 2 analyses the major goals that may underlie communicative interaction of an organization with a larger audience. We identify yield, brand, and reputation management as three major aspects around which communication may be centered. They can be identified as variations of general value management differing in their short-versus-long-term orientation towards commercial goals, as well as in their overall connection to financial value orientation. Section 3 sketches the major technical elements that we developed to implement a common value management framework. Section 4 discusses some of the related work and directions for future research activities. Finally, conclusions are provided in Section 5.

2 The Aspects of Value Management

Scalable, multi-channel communication is a difficult challenge. In order to better understand it, we want to clarify the various underlying goals that it should achieve. Agents often connect (directly or indirectly) economic interests with their communication activities. In the following, we discuss different economic contexts for the communication approaches of organizations.

2.1 Yield Management

Yield or revenue management "is an economic discipline appropriate to many service industries in which market segment pricing is combined with statistical analysis to expand the market for the service and increase the revenue per unit of available capacity" [IDEAS, 2005].[1] Short-term increase of income is a valid target for a business entity; however, it is quite tricky to realize in a multichannel world. For example, hotels are confronted with a multitude of on-line booking channels. Hotels should provide their available rooms and rates to most, if not all of these channels to prevent missing their potential customers. For many channels, visibility is achieved through low prices. However, channels also often require price constraints on the price offers of other channels. Some channels generate costs without guaranteeing actual income. Let us discuss some important issues: A hotel currently provides active booking facilities on its hotel website, through booking.com, expedia.com, Facebook, and Google Maps. It wants to increase the overall conversion rate for bookings in all of these channels, and therefore needs answers to the following questions:

- How is my price positioned towards the price of my competitors?

[1] http://en.wikipedia.org/wiki/Yield_management, and Revenue_management

- What is my reputation in all of these channels? More than 90% of all Internet users are already reading product reviews and more than 50% have indicated that their purchasing decisions are based primarily upon them.
- Am I adequately represented in all these channels (comments, reviews, etc.)?
- The hotel could reduce its price in channel X to maintain visibility and to increase bookings through this channel. Due to legal constraints, it has to reduce its price in many other channels which leads to reduced revenues from those bookings.
- A hotel needs recommendations for what needs to be done and the support to do it, e.g. possible actions would be to reduce their price by 10% or to include more amenities and supplements, to be more active on Facebook in order to increase social media links and conversations, or if guests complain about the coffee it could therefore be alerted to improve this service.
- A hotel wants to announce rooms through Google Maps. It therefore starts to pay for each click through this search interface. If many clicks fail to lead to a booking, the hotel may begin to lose a significant amount of money.
- The same scenario could occur with static on-line coupons[2] that offer a 50% discount through a coupon platform that requests an additional 25% for each coupon. This can easily end in negative revenue and lacks dynamicity. Imagine a "*magic*" dissemination button for a bar owner that can announce dynamically a happy hour, special offers, interesting news etc. on-the-fly to the right circle of interested public, establishing the bar as a trendy place where a hip crowd is getting together.

Many solutions to yield management are based on complex statistical methods and complex domain assumptions on how variation of the price can influence the number of bookings of a service. However, a multi-directional multi-channel approach must also rely on *Swarm intelligence*[3]. Observing in real time the reaction of customers and competitors will be key to achieving on-line marketing. Adopting your offer and your price dynamically in response to the behavior of your (on-line visible) environment will become critical to economic success.

2.2 Brand Management

Yield management tries to maximize the immediate revenue of an organization. However, communication is also very important in relation to the long-term value of a company. Actually, the reputation of a company can be viewed as one of its most important assets. Proper management, such as managing the value of **brands**,[4] may be essential for its long-term economic success. This may conflict with revenue

---

[2] http://en.wikipedia.org/wiki/Coupon

[3] http://en.wikipedia.org/wiki/Swarm_intelligence

[4] "The American Marketing Association defines a brand as a "name, term, design, symbol, or any other feature that identifies one seller's good or service as distinct from those of other sellers." http://en.wikipedia.org/wiki/Brand

management. In many cases, it may be useful for short-term income management to reduce the price of the offering, which on the other hand can diminish and undermine the long-term income that is generated through a general price profile indicating quality and exclusivity.

2.3 Reputation Management

The economic impact of proper reputation management is evident when we talk about the reputation of economic entities. However, non-profit organizations also have a need for general *reputation management* and *public campaigns*. [5] "Reputation is the opinion (more technically, a social evaluation) of a group of entities toward a person, a group of people, or an organization on a certain criterion. It is an important factor in many fields, such as education, business, online communities or social status."[6] Here, it is not the direct and intermediate economic income that matters. It is rather about maintaining or increasing the appreciation an organization, topic, or certain approach gains in the public eye. However, even a campaign on a public issue has an immediate economic dimension to it: trying to use the available budget for it in the most effective way. Therefore, providing means to increase the effectiveness and efficiency of public campaigns is of high value.[7]

2.4 Value Management

All of the issues above could be viewed as facets of Value Management, where value is defined as *the regard that something is held to deserve, i.e., its importance*. Online, multi-channel and bi-directional Value Management is about disseminating [8], communicating, and interacting with large, on-line communities to increase the value of a certain entity or issue. The value managed could cover issues such as importance, economic short-term income, or long-term value. [Kasper et al., 2010] identifies the following activities as part of an on-line based value management: Reputation management; Competitive Intelligence, i.e., Competitor Observation; Market Analysis; Influencer Detection; Trend Analysis; Market Analysis; Crisis Management; Issue Management; Campaign Monitoring; Product and Innovation Management; Customer Relationship Management; Risk Management; and Event Detection. Obviously, these activities overlap and share many common elements. It

---

[5] E.g. http://www.readwriteweb.com/archives/how_to_manage_your_online_reputation.php.

[6] http://en.wikipedia.org/wiki/Reputation

[7] Also, in the case of political parties, the number of votes they collect can be seen as their "economic" value.

[8] "To disseminate (from lat. disseminare „scattering seeds"), in terms of the field of communication, means to broadcast a message to the public without direct feedback from the audience." http://en.wikipedia.org/wiki/Dissemination

would be interesting to reduce these activities to the set of atomic tasks from which they are composed.

3 A Methodological Approach towards Common Value Management

We start this section by introducing the underlying idea and major structure of our approach. We then discuss our role model, tool support, and finally, we sketch some applications of our approach.

3.1 Separating Content and Channel to enable various dimensions of reuse in transactional communication

"I am about to propose the existence of something called the knowledge level, within which knowledge is to be defined." [Newell, 1982]

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.[9] So one has:

- *information models*, that define the type of information items in a domain;
- a *channel model* (or communication model), that describes the various channels, the interaction pattern, and their target groups;
- *mappings* of information items to channels through weavers; and finally,
- a library of *implemented wrappers* for actual channel instances.

What is essential is to *distinguish* the communication or channel model from the conceptual descriptions of the information.[10] Our approach requires the creation of a communication model (i.e., an increasingly complete model of channels), and knowledge models for each vertical (such as research projects, research institutes, associations, hotels, restaurants, tourist events, medical doctors, etc.), and finally linking the knowledge model with the communication model through a weaver that weaves concepts with channels. This will *not* be cheap! However, you pay one price for an entire vertical, i.e. for thousands or up to millions of business units. Even if it would cost the same amount as doing it manually directly on a channel basis for ten of them, it is obviously cheaper if the number of business units is higher than ten. And on the contrary, we think that even if done directly for one business unit, it quickly turns into profit by saving on manual dissemination work through a mechanized communication model. After all, data and information can be expressed at the conceptual level, which the domain expert understands. Mapping of the different

[9] See also as an excellent presentation on this idea: http://www.slideshare.net/reduxd/beyond-the-polar-bear

[10] In analogy to style sheets that separate the contents from its presentation.

communication means (where he is not at all an expert) is done automatically after the first implementation. The difficult dissemination through channels is done automatically through proper channels that are attached to these concepts.

Currently, all commercially available solutions are only channel centric and do not provide any built-in support for what needs to be disseminated or where to disseminate what piece. In our approach, a knowledge-model is built and explicitly linked with the channel model. This must be done once for a hotel, and can then be reused for millions of them. That is, we aim for the major elements of reusability:

1. The same information element can be *reused* for various channels through its channel independent formulation using the information model.
2. The information model is developed as domain ontology for a certain vertical area such as tourist accommodations, gastronomy, medical doctors etc. Therefore, it can be *reused* for various agents active in the same vertical domain.

These elements of reusability deliver the major contribution to the scalability of our approach.

3.1.1 Information Model

An information model is an ontology that describes the information items that are used in typical communication acts in a certain domain. Many methodologies for building such ontologies have been developed, compare [Staab & Studer (eds.), 2009]. Building ontologies can be a time-consuming and expensive process. Fortunately, we have a strong modeling bias that helps us to significantly guide and therefore reduce such an effort. We do not aim to model a domain as such, through a very deep model that allows arbitrary (transactional) applications. We can rather focus on the major and typical information items that are used in the on-line dissemination and communication processes. Therefore, the size of these ontologies in our case studies (see Section 3.4.), were moderate (around 100 concepts and properties), and many of these concepts and properties could be reused between different use cases. As a result, there was a reduced effort in building informal domain models (less than one person month). After defining an informal model, we formalized this ontology (see [Bauereiß et al., 2011] for more details) in a simple sublanguage of OWL-2, since we foresee little need for reasoning about it. We model structured information items as concepts and non-structured ones as properties, i.e., we assume simple non-structured values for properties. Actually, we separated a generic ontology common to all three use cases and specific refinements. The classes of the core ontology are depicted in Figure 1 and its properties in Figure 2.

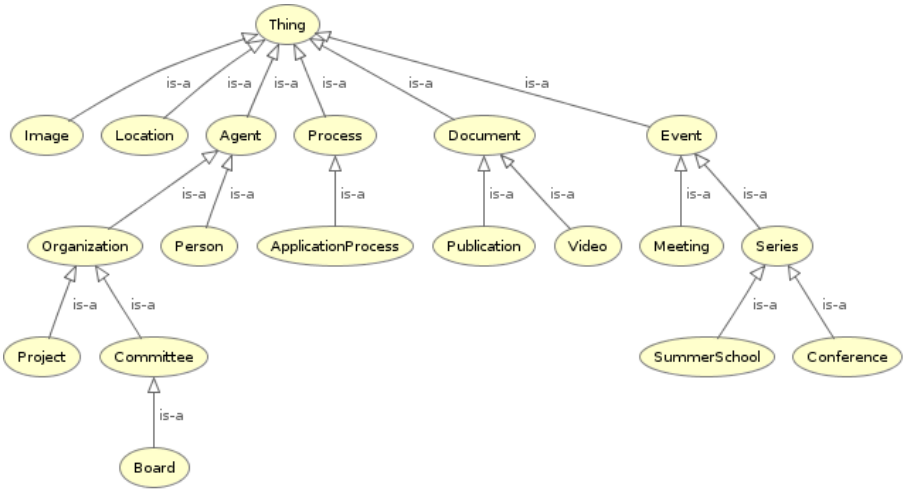

Figure 1. Classes of the core ontology

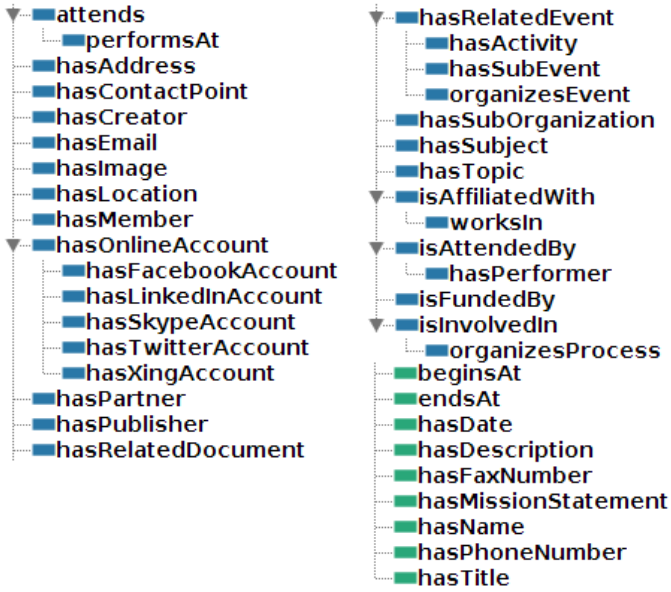

Figure 2. Properties of the core ontology

In general, our models are rather simplistic, but this reflects the fact that we do ***not*** model a domain with all its complexity but rather the information chunks that are disseminated about it. That is, our modeling perspective and grainsize are defined by the consideration that an information provider can make informal sense of a chunk and that this chunk can be mapped onto a communication channel.

In an intermediate phase of our journey, we also tried to directly use some LOD vocabularies to model these ontologies.[11] However, we took some important lessons from this enterprise:

- Our domain models were suddenly unintelligible to the domain experts. The LOD vocabularies used different terms and made different and non-intuitive modeling decisions compared with our informal domain models (especially not the ones in the eTourim case study).
- It was extremely hard to decide which term to take from which vocabulary. The terminologies were either redundant or terms had different but overlapping coverage.
- Suddenly, we had to deal with a large number of properties for which we had never asked.

"In contrast to the heterogeneity of the Web, it is beneficial in the application context to have all data describing one class of entities being represented using the same vocabulary … it is thus advisable to translate data to a single target vocabulary". [Schultz et al., 2011] We draw an important conclusion from this: *For us, LOD vocabularies are not means to describe our content models, i.e. they were not really useful for deriving domain models.* We decided it would be better to *interpret them as channels*. That is, we model our information items in a Domain Ontology that is understandable by the domain experts. Interaction with them is essential to our approach and therefore understandability of our means towards these domain and communication experts. We then provide mappings (though our weaver, see Section 3.1.3) that export and/or import information to or from terms of various LOD vocabularies. Whenever we see a significant uptake of a vocabulary by a target group that we want to talk and disseminate to, we establish such a link. In the end, a term in a LOD vocabulary is treated similarly to a URI from our web pages. We export or import some of our content to or from it. *For us, LOD vocabularies are means to disseminate and share information and not means to model information*. Ontologies are always on the brink of being a very specific and well-defined domain model derived from certain first principles, being very useful for a specific purpose in contrast to broadly used and consensually developed models used for sharing information between different viewpoints. Consequently, we live in a world of multiple ontologies. "We no longer talk about a single ontology, but rather about a network of ontologies. Links must be defined between these ontologies and this network must allow overlapping ontologies with conflicting – and even contradictory – conceptualizations." [Fensel, 2001] We achieve this by weaving our models with

[11] We had used a mélange of Dublin Core, FOAF, schema.org, and GoodRelations.

LOD vocabularies when we see a gain in broadening our range of communication through them.[12]

3.1.2 Channel Model

> *"A growing number of retailers are becoming increasingly multichannel as more of their sales are coming through their web divisions than ever before."*
> *[Mulpuru et al., 2011]*

"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." [13] In on-line 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

---

[12] "The way the Semantic Web works, and this is what makes it very different from everything else, is that you use a mixture of global ontologies like foaf:Person and dc:title and a number of other ontologies which are relevant, and then add on some more to make up what you need. If this sounds like a mess …" Tim Berners-Lee, email communication, Mon, 20 Feb 2012. Actually, we move the handling of this "*mess*" in the mapping of various vocabularies and free the user working at the information model level from it.

[13] http://en.wikipedia.org/wiki/Communication_channel

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.[14]

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

[14] [Finzen et al., 2010] propose a slightly different categorization of channels, however, call them different types of information.

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[15]. 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. [16] *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, such as that shown in Figure 3, which can be analyzed by the format and vocabulary used. "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.[17] 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

[15] Write access cannot be provided due to spamming.
[16] http://www.nytimes.com/
[17] See [Tennison, 2012] for more details.

integrates best with the W3C meta data 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. Figure 4 shows that the use of RDFa has increased rapidly, whereas the deployment of microformats in the same period has not advanced remarkably. Consequently, we are focusing on RDFa and Microdata in our dissemination approach.

Figure 3. Structured content presentation[18]

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. [19] Currently, we use Dublin Core, FOAF, GoodRelations, and schema.org.

Notice that we use each term of a vocabulary as a potential dissemination channel. For example, for the PlanetData fact sheet we publish pieces of the information using the following vocabulary terms: schema:url, foaf:topic, dc:creator, dc:date, dc:subject, and dc:title.

[18] Taken from http://www.google.com/support/webmasters/bin/answer.py?answer=99170
[19] More than a hundred of them are listed at http://labs.mondeca.com/dataset/lov/index.html.

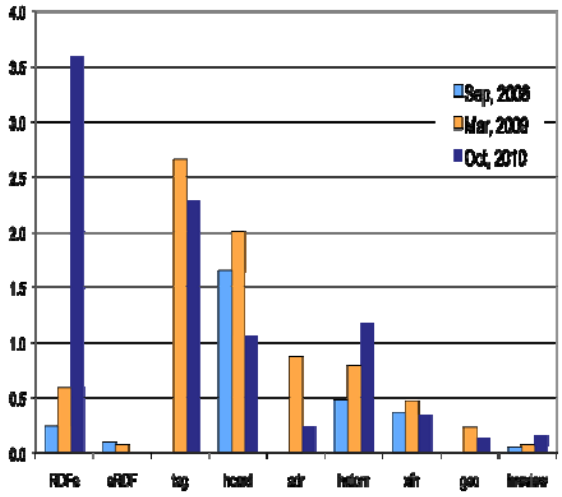

Figure 4. Microformats and RDFa deployment on the Web (% of all web pages)[20]

3.1.3 Weaver

The central element of our approach is the separation of content and communication channels. This allows reuse of the same content for various dissemination means. Through this reuse, we want to achieve scalability of multi-channel communication. The explicit modeling of content independent from specific channels also adds a second element of reuse: Similar agents (i.e., organizations active in the same domain) can reuse significant parts of such an information model.

Separating content from channels also requires the explicit alignment of both. This is achieved through a weaver. Formally, a weaver is a set of tuples of nine elements:

1. An *information item*: As discussed in Section 2, it defines an information category that should be disseminated through various channels.
2. An *editor*: The editor defines the agent that is responsible for providing the content of an information item.
3. An *editor interaction protocol*: This defines the interaction protocol governing how an editor collects the content.

Elements 1 to 3 are about the content. They define the actual categories, the agent responsible for them, and the process of interacting with this agent. Elements 4 to 9 are about the dissemination of these items.

[20] http://tripletalk.wordpress.com/2011/01/25/rdfa-deployment-across-the-web/

4. An *information type*: We make a distinction between three types of content: an instance of a concept, a set of instances of a concept (i.e., an extensional definition of the concept), and a concept description (i.e., an intensional definition of a concept).
5. A *processing rule*: These rules govern how the content is processed to fit a channel. Often only a subset of the overall information item fits a certain channel.[21]
6. A *channel*: The media that is used to disseminate the information.
7. *Scheduling information*: Information on how often and in which intervals the dissemination will be performed which includes temporal constrains over multi-channel disseminations.
8. An *executor*: It determines which agent or process is performing the update of a channel. Such an agent can be a human or a software solution.
9. An *executor interaction protocol*: It governs the interaction protocol defining how an executer receives its content.

First, the information types distinguish whether one wants to disseminate a general description of the information item, an instance of the information item, or a set of all instances. For example, we want to find an overall description of scientific presentations (what is their general theme) and a set of all presentations at a defined place on the web. The former may be placed on the project website and the later may be placed on SlideShare as a means to share presentations. Finally, a single instance may be broadcast as news through the various news broadcasting channels. Now, take a single presentation as an example. The title, author, abstract, and event it was given may form the news. The title, author, and a short notion of the event may define a tweet, and the slides themselves may go to SlideShare. That is, the information item must be processed to fit the various dissemination channels. A channel is a URI or an API of an existing web service. Scheduling information defines temporal constraints for dissemination in a single channel and for dependencies between multi-channel dissemination. For example, a new presentation will be announced once. However, an event may be announced as soon at it is defined and a reminder may be sent out when certain deadlines (for submitting papers or for early registrations) are near. News may first be published on the website. Then, an excerpt of the news together with its URI will be published as a tweet.

A weaver is basically a large collection of tables specifying what is disseminated by whom to where. Interaction protocols, rules, and constraints further guide this process. Such a manual is of extreme importance to manage the on-line communication process. Obviously, it determines the need to implement and mechanize essential aspects of it, improving its scalability. However, a major step is to structure the process towards a mechanizable routine.

---

[21] In case of LOD this can be an R2R mapping rule [Bizer & Schultz, 2010].

3.1.4 A Spiral as Process Model

The basic SMCR model of communication[22] is unidirectional. A sender sends a message through a channel to a receiver. The direction of the communication and the different roles are fixed. Actual communication is more complex. Agents interact and communicate in parallel, permanently alternating their role in these acts of communication. Therefore, we have adopted the *transactional model* of communication and its underlying premise that individuals are simultaneously engaging in sending and receiving messages (cf. [Barlund, 2008]).[23]

Consequently, our approach not only disseminates information, but also deals with the aggregation of feedback and impact by simply going through the dissemination chain in the opposite direction, collecting responses in the various channels and integrating them under the appropriate knowledge item. We not only talk, we also listen to responses. And we do not get these responses scattered over multiple channels. Instead, they are aggregated and presented at the level of the domain specific concept to which they refer. Finally, our approach can also aggregate information from channels without prior publication activity. Communication situations initiated by an external agent can be handled through standing queries over pre-defined channels and dynamic web screening.

Therefore, a holistic methodology for supporting communication must support the following subtasks that basically form a circle or spiral (see Figure 5):

- Design of an information item
- Dissemination of the information item using suitable channels and places
- Observation of communication acts
- Measure, analysis, and aggregation of the information published

These activities form a circle that we call the *life cycle model of communication*. Reactive communication starts with the observation task; active communication starts with the design phase. In any case, when started, one has entered an-in principle-infinite loop. Obviously, these tasks can and should be parallelized once initialized.

[Li & Bernhoff, 2008] identify five levels in reputation management:

- *Listening*: learn about the topic and the community;
- *Talking*: participate in discussions;
- *Energizing*: let other key players talk for you;
- *Supporting*: support the community through initiatives and platforms; and
- *Embracing***:** the on-line community is used to further developing the company value in a crowd sourcing approach.[24]

[22] http://en.wikipedia.org/wiki/Models_of_communication

[23] Or in Web2.0 terms, users are *prosumers*, i.e., consumer and producer of information.

[24] See also [Helbing & Konitzer, 2010] for a similar and refined model.

Each of these levels requires a different instantiation of our process model. In general, both models reflect the bi-directional character of communication in extension of more simplistic SMCR models.

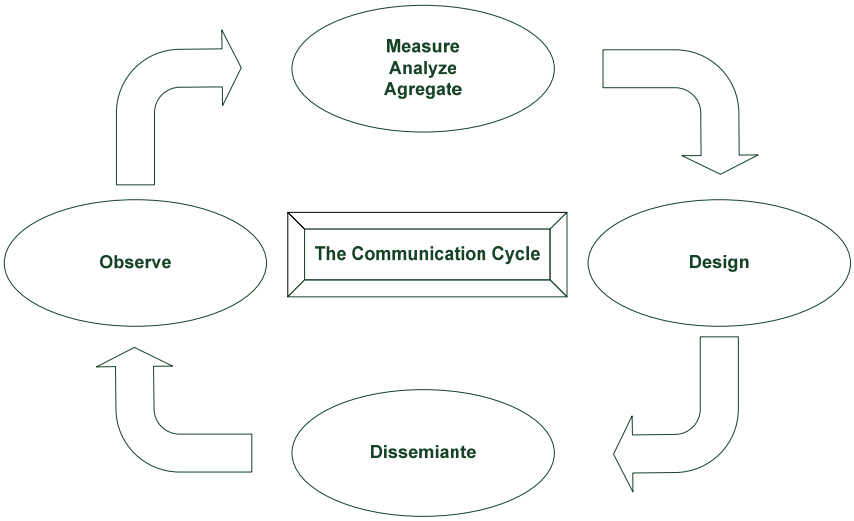

Figure 5. The Life Cycle of Communication

3.2 Role Model

Editors are assigned to information items, responsible for producing or collecting their information instances. In general, an editor can also be the executor, directly publishing the information. However, expertise in a certain information domain may not necessarily correspond to technical expertise and even if it does, it may not be a very efficient way of distributing labor. Only if a fully automated and easy to use software solution is provided can this model make sense. Otherwise, a person with more technical skills often helps in disseminating the information. Again, an interaction protocol has to be defined for interacting with this person. Recursively, some of his tasks may be manual (importing contents into a content management system such as Drupal) and some can be fully mechanized (like producing a feed and a tweet automatically for certain information items introduced into Drupal). We identified five different roles involved in this process (see Figure 6):

- The *communication manager* that actively reads and writes information in the multi-channel space and manages the overall communication process.
- The *quality manager* that routinely checks the outcome of the process and the impact that is achieved through it.

- The *editors* that provide information that should be disseminated or that infer actives from information provided by others.
- The *web manager (executor)* is an expert in web technology who is able to publish information with current web technology including content management systems such as Drupal, email lists and Web 2.0 services such as Twitter, Blogs, RSS, and has the means to share information, cooperate, or organize communities through SNS sites.
- The *repository manager (executor)* is an expert in semantic web technology in terms of syntax, implementations via repositories, and various vocabularies used to publish this information. In a nutshell, the web manager manages the web of documents and the repository manager manages the web of data.

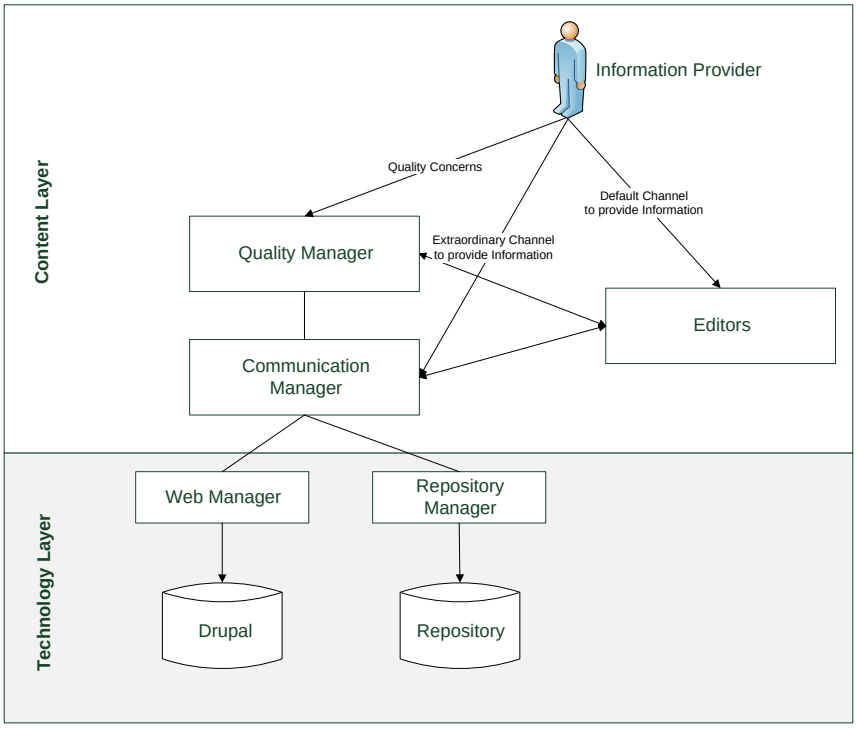

Figure 6. Roles

3.3 Technologies

Obviously there is an important need for methods and integrated tools that cover the multi-channel bi-directional aspects of value management and provide highly scalable

and effective solutions. We are developing a communication platform and methodology for providing this based on Semantic Technologies, Human Computation, and Social Media Analysis.

3.3.1 Semantic Technologies

Semantic Technologies (cf. [Domingue et al (eds.)., 2011]) are a stream of research combining web technology, artificial intelligence, natural language processing, information extraction, database technology, and communication theory for empowering computers to provide better support for processing, combining, and reusing information represented as structured and unstructured data.

We use content management tools such as Drupal 7 to include RDFa, microdata, and microformats in the web documents. The data will also be exported from Drupal into OWLIM[25] to support direct RDF[26] and SPARQL (see Figure 5).

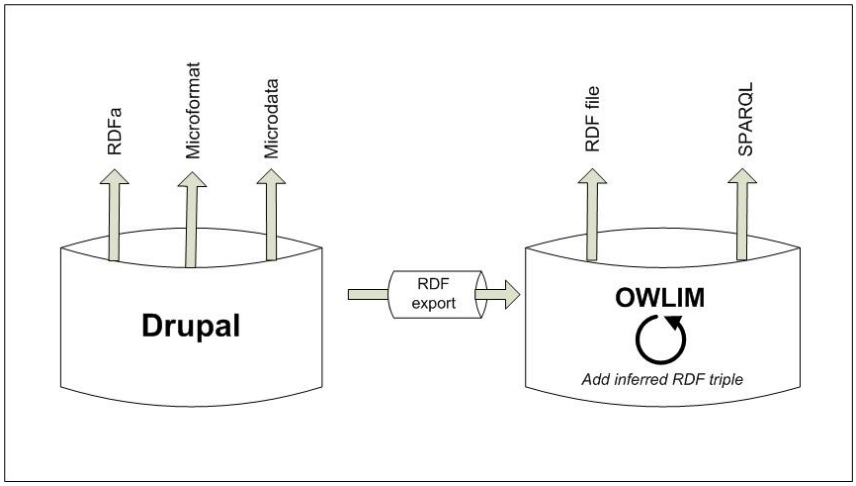

Figure 5. Technical means to publish Semantic Data

Either the editors, or alternatively, the dissemination manger enrich the content for on-line presentation by adding links and tags to the presented information. For this purpose, tools such as the following are used:

[25] "OWLIM is a family of semantic repositories, or RDF database management systems, with the following characteristics: native RDF engines, implemented in Java and compliant with Sesame, robust support for the semantics of RDFS, OWL Horst and OWL 2 RL, best scalability, loading and query evaluation performance" http://www.ontotext.com/owlim

[26] The RDF file hast to be generated by OWLIM to include inferred triples.

- *KIM*[27] offers the ability to create semantic links between your documents, data, domain models, and linked data; find mentions of entities, relationships, and facts in texts; and search and navigate your information space in multiple ways.
- The *OpenCalais Web Service*[28] creates metadata for the content you submit. Using natural language processing (NLP), machine learning and other methods, Calais analyzes your document and finds the entities within it. Additionally, Calais returns the facts and events hidden within your text. The metadata gives you the ability to build maps linking documents to people, companies, places, products, events, geographies, etc. You can use those maps to improve site navigation, provide contextual syndication, tag and organize your content, create structured folksonomies, filter and de-duplicate news feeds, or analyze content to see if it contains information you care about.
- *Zemanta*[29] is a platform for accelerating on-line content production for any web user. It searches the web for the most relevant images, smart links, keywords and text, instantly serving these results to the user to enrich and inform their content.

3.3.2 Human Computation

Human Computation (cf. [Ahn, 2005]) is a research field that aims to integrate human intelligence and human cognition into the computer-managed, automated execution of tasks that are beyond the power of the state of the art of pure algorithmic approaches. Until artificial intelligence has successfully enabled computers with sufficient human-like intelligence, on-line communication will continue to require and involve human labor. Understanding text, generating useful output, and interacting properly cannot be achieved without having a human in the loop. However, as we also saw it will not scale if too many tasks are left to humans. Therefore, we need to develop an architecture that tries to maximize the amount of tasks that are mechanized and that provides smooth integration of humans for tasks that cannot be fully automized. Such architecture may also have the potential to be applied in other domains with similar characterizations. This can include crowd-sourcing[30] initiatives where we develop methods to define incentives for large user communities to provide content needed for your communication strategy through a distributed community effort.[31]

[27] http://www.ontotext.com/kim
[28] http://www.opencalais.com/
[29] http://www.zemanta.com/
[30] http://en.wikipedia.org/wiki/Crowd_sourcing
[31] For example, see the results of the Insemtives project at http://www.insemtives.eu/ or the social stock market *empire avenue* http://empireavenue.com/about/

3.3.3 Social Media Analysis

Social Media (cf. [Kaplan & Haenlein, 2010]) is a term used mostly for web-based techniques of human-to-human communication that stresses the social, topical, and contextual relations between communicating individuals, allowing real-time interaction with a large, yet specific audience of partners. Social Network Analysis emerged from the area of sociology that studies specific social phenomena. More recently, it was carried out under the umbrella term of complex network analysis, a field that studies properties in large, complex graphs. It became increasingly more popular through the huge success of social networks such as Facebook, Twitter, Flickr and others (cf. [Bonchi et al., 2011]). Web mining [Kosala & Blockeel, 2000] is the use of data mining techniques to automatically discover and extract information from web documents and services. Machine learning, a branch of artificial intelligence, is a scientific discipline concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases.

For reputation management, we require support for impact analysis, multi-triggered bi-directional communication, and standing queries over streams. We enumerate some tools for this[32,33]:

- *Appinions*[34], *Empire Avenue*[35], *PeerIndexes*[36], *Proskore*[37], and *TRAACKR*[38] identify, rank, and score on-line "authorities".
- *Boardreader*[39] is a search engine for Forums and Boards.
- The *ComScore Media Metrix Suite*[40] project is a framework for web traffic measurement that is able to provide traffic statistics along several properties (age, region of users, etc.). Therefore, Media Metrix uses a representative set of Internet users, a weighting algorithm, and enumeration surveys to statistically make a projection of the whole population of Internet users. Additionally, ComScore introduced other tools such as mobile Internet traffic measurement and advertisement impact analysis software.

---

[32] In addition to tools there are also companies that offer reputation management as a service, f.e., www.reputation.com.

[33] See also [Kasper et al., 2010], [Solis, 2012], [Stavrakantonakis et al., submitted], and http://www.somemo.at/?p=474.

[34] http://appinions.com/

[35] http://empireavenue.com/

[36] http://www.peerindex.com/

[37] http://proskore.com/

[38] http://www.traackr.com/

[39] http://www.boardreader.com/

[40] http://www.comscore.com/Products_Services/Product_Index/Media_Metrix_Suite/Media_Metrix_Core_Reports

- *Facebook Insights*[41] provides Facebook Page owners and Facebook Platform developers with metrics, just as *CircleCount*[42] does for Google+. *Google+ Ripples*[43] graphically illustrate the sharing of posts in Google+.
- *Google Alerts*[44] are email updates of the latest relevant Google results (web, news, etc.) based on your choice of query or topic.
- *Google Analytics*[45] and *Yahoo! Web Analytics*[46] are web analytics solution providing insight into website traffic and marketing effectiveness. F.e., Google Analytics provides analysis and optimization tools such as Urchin[47], Website Optimizer, Webmaster Tool[48], Insights for Search[49], and further tools for improving advertisement and Search Based Keyword Tools included in your website. *AWStats*[50], *Open Web Analytics (OWA)*[51] and *Piwik*[52] are open source web analytics software with similar functionalities.[53]
- *Google Trends*[54] analyzes a portion of Google web searches to compute the number of searches that have been done for the terms you enter, relative to the total number of searches done on Google over time.
- *ifttt*[55] provides a multi-channel trigger and action engine that allows automatic reaction to events in the information space by observing a multitude of channels and executing guarded transaction rules acting on these channels.
- *Klout*[56] and *Kred*[57] measures influence and impact of communication acts on social media.
- *Open Status Search*[58] finds out what people on Facebook are talking about in public.

---

[41] http://www.facebook.com/help/search/?q=insights
[42] http://www.circlecount.com/
[43] http://www.google.com/support/plus/bin/answer.py?answer=1713320
[44] http://www.google.com/alerts/
[45] http://www.google.com/analytics/
[46] http://web.analytics.yahoo.com/
[47] Urchin helps website owners better understand their online marketing initiatives, website traffic characteristics, and visitors browsing experience.
[48] Google Webmaster Tools provides you with detailed reports about your pages' visibility on Google.
[49] With Google Insights for Search, you can compare search volume patterns across specific regions, categories, time frames and properties.
[50] http://awstats.sourceforge.net/
[51] http://www.openwebanalytics.com/
[52] http://de.piwik.org/
[53] Products in this area are *Compete*, *Coremetrics*, *Omniture*, Optimizely, and WebTrends.
[54] http://www.google.com/trends
[55] http://ifttt.com/
[56] http://klout.com/home
[57] http://kred.com/
[58] http://www.openstatussearch.com/

- The tool suites of *eCairn*[59], *MediaMetrics*[60], and *Radian6*[61] allow to seek out all conversations surrounding a topic on the social web, measure and summarize them, and to react and publish at social media channels.
- The *Social Media Dashboard*[62] manages multiple social profiles, allows scheduling messages and tweets, can track brand mentions, and analyzes social media traffic.
- *Social Mention*[63] is a social media search and analysis platform that aggregates user generated content into a single stream of information. Social Mention monitors 100+ social media properties directly including: Twitter, Facebook, FriendFeed, YouTube, Digg, Google etc.
- *Technorati*[64] helps bloggers succeed by collecting, highlighting, and distributing the global online conversation. Founded as a blog search engine and directory, Technorati indexes more than a million blogs. Technorati tracks not only the authority and influence of blogs, but also the most comprehensive and current index of who and what is most popular in the Blogosphere.
- *Twazzup*[65] operates a real-time news platform.
- *TweetDeck*[66] provides multi-channel integration for reading and writing for Twitter, Facebook, MySpace, LinkedIn, Foursquare and Google Buzz
- *Twibes*[67], *Twitalyzer*[68], *TweetLevel*[69], TweetReach[70], TwitterGrader[71] provide means to analyze content on Twitter.

3.4 Use Cases

We developed and applied our approach in three major case studies: the European Semantic Web Conference Series (ESWC)[72], the PlanetData project[73] and the Semantic Technology Institute (STI) International research association[74].

[59] http://ecairn.com/
[60] http://www.media-metrics.de/
[61] http://www.radian6.com/
[62] http://hootsuite.com/
[63] http://www.socialmention.com/
[64] http://www.technorati.com/
[65] http://www.twazzup.com/
[66] http://www.tweetdeck.com/
[67] http://www.twibes.com/
[68] http://www.twitalyzer.com/
[69] http://tweetlevel.edelman.com/
[70] http://tweetreach.com/
[71] http://tweet.grader.com/
[72] http://eswc-conferences.org/
[73] http://www.planet-data.eu/
[74] http://www.sti2.org/

- The mission of the *Extended Semantic Web Conference (ESWC) series* is to bring together researchers and practitioners dealing with different aspects of semantics on the web. Founded in 2004, the ESWC builds on the success of the former European Semantic Web Conference series, but seeks to extend its focus by engaging with other communities within and outside ICT, in which semantics can play an important role.
- *PlanetData* is a semantic technology project funded by the European Commission. It aims to create a durable community made up of academic and industrial partners working on large-scale data management.
- *STI International* is a global network engaging in research, education, innovation and commercialization activities on semantic technologies working to facilitate their use and applicability within industries and society as a whole. STI International is organized as a collaborative association of interested scientific, commercial and governmental parties that share a common vision.

Around 80% of the information items of ESWC, PlanetData, and STI International are interchangeable due to some simple renaming (e.g., core and associate partner versus partner and member). This is excellent news and a hint for scalability especially given the fact that we talk about a research *project* and a research *association*. This could imply that an even higher degree of reuse could be achieved when applying our information model to tens of thousands of European research projects (and hundreds of thousands of research projects or millions of projects) on the one hand, and millions of associations on the other. This is actually the second major assumption of our approach.[75] Reuse of the information model in a certain vertical area. The costs to build an information models are quickly paid back when applicable to several entities in a domain. In the end, this is the SAP business model applied to on-line communication. These models empower simple non-IT users to communicate at the level of their domain knowledge rather than at the symbol level of various channels and these models can be reused between different players in the same vertical.

Based on our approach ESWC, PlanetData, and STI International are now managing their on-line appearance. In total, we have identified around *five hundred* different semantic and non-semantic channels in these case studies that are used to disseminate elements of the information model. Obviously, such a bandwidth requires a structured and mechanized approach. Based on our approach, around 300 concepts and properties, 500 channels, i.e., more than 100,000 potential content-to-channel mappings are run efficiently by a very small dissemination team.

[75] The first one is that is that it will pay back to model the information independent from the multitude of dissemination channels, ensuring reuse over them.

3.5 Our proposed solution in a nutshell

The core idea is to introduce a domain-specific, channel-independent model that explicitly separates content from channel. The next step is to once again intelligently interweave the content with the channels. Our approach of modeling communication, communication channels and target groups inherently bears the advantage of uniformly accessing the provided data and thereby allowing the processing of data that is beyond state of the art. For example, yield management could be realized utilizing reputation and usage values collected from different channels. Furthermore, the abstraction layer allows multi-channel communication. Human computation could increase the process where automated algorithms lack efficiency, for example the translation of communicated content into other languages. Combining these different areas of technology provides a long-term roadmap for research, engineering, and commercial exploitation.

4 Related and Future Work

4.1 Related Work

Many aspects of our work clearly relate to different fields that have been explored before. Generally, we see two specifically related areas: *Ontology-based content management systems (CMSs) for websites* and *Semantic matchmaking of senders and receivers of content*. Both areas will be briefly described and compared.

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. This is related to our weaver concept introduced in Section 4, but, in contrast, the weaver adapts to different channels rather than to different information sources. A year later, in [Sheth et al., 2002], Sheth et al. introduce 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. [Garcia et al. 2008] introduce “The Rhizomer Semantic Content Management System” which integrates services with metadata browsing, editing, and uploading, continuing their earlier work on the Knowledge Web portal. [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[76] 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 straight-forward. In contrast, our system accounts for different information needs of various and heterogeneous channels and therefore enables the distribution of content through different portals. Finally, the European project Interactive Knowledge Stack (IKS)[77] focuses on porting semantic technologies to CMS software solutions.

In a nutshell, all these approaches aim either to help the user publish semantic data or to use semantic methods to support the content management process for maintaining websites. We are taking these approaches and generalizing them to support the overall management of content dissemination in a multi-channel and bi-directional communication setting. Further, we augment the technical approach with a methodology and the approach of using vertical domain models, which are shared and reused in a vertical area instead of being used for a single application only.

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] present 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 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 performs 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 channel level to the content level, but we see the need for human involvement in defining the content-channel mapping and at the content level. However, as we expand towards a

[76] http://www.bbc.co.uk/worldcup
[77] http://iks-project.eu/

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] and [Stavrakantonakis et al., submitted] [78] lists a large number of them that cover parts of these tasks. However, there is an important need for methods and integrated tools that cover the multi-channel bi-directional aspects of value management and provide highly scalable and effective solutions. Obviously, the goal to develop a Common Value Management Framework (CVMF) based on combining these different areas of technology provides a long term roadmap for research, engineering, and commercial exploitation.

4.2 Future Work

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

- *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 of validating whether or not the recipient received or understood the message. It is also often preferable to have a full-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 wish to address. So far in our channel model, we do not distinguish between different target groups in different channels. However, different target groups reside in 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 - converting an information item into different formats or automatable transformations such as extracting images, videos or extracting and shortening web links from pieces of content, and transforming multimedia content into a different format. Both of these problems can be commonly solved, however, adapting content in a way that requires creativity and human intelligence is still a challenging

[78] See also [Solis, 2012] and http://www.somemo.at/?p=474.

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.

- *Crowd sourcing*. Crowd sourcing is an online approach of outsourcing tasks that computers cannot solve, to humans. There are different incentives for people to work on crowd sourcing problems. The most common one is small amounts of money (e.g. Amazon Mechanical Turk[79], Clickworker[80] or Zhubajie[81]) but other sources of motivation are common such as competition (e.g. TopCoder[82]) or language learning (e.g. DuoLinguo [83]). Crowd sourcing will be used to smartly enhance sentiment analysis, natural language processing, and translating algorithms with human intelligence where needed until artificial intelligence has reached a more satisfactory level. CrowdDB [Franklin et al., 2011] is a successful example of combing the two computation paradigms.
- *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 (cf. [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 they 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 social values*. The success of online social networks leads to an enormous amount of data that may be analyzed in order to find out about social structures and relationships. A mature way of quantifying values like popularity, authority, influence or reach has yet to be found. Various tools such as PeerIndex[84] or Klout [85] already attempt to measure it, but consider only very small parts of the whole social media landscape. Also, there are other attempts to measure various factors in social media using different

[79]https://www.mturk.com/mturk/welcome
[80]http://clickworker.com/
[81]http://www.crowdsourcing.org/site/zhubajie/wwwzhubajiecom/2118
[82]http://www.topcoder.com/
[83]http://duolingo.com/
[84]http://www.peerindex.com/
[85]http://klout.com/home

techniques or different forms of proximity to forecast activity (cf. [Lermann et al., 2011]).

- *Quantification of brand and reputation*. Similar to the quantification of social values, brand and reputation have to get a countable unit as well. Likewise, it will be very challenging to find fitting metrics, since already existing measures such as brand equity are considered meaningful by a small share of marketing professionals (cf. [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 on different parameters, such as pricing, capacity, and demand (cf. [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 service/product could be less than in other channels where it is well known. On the other hand, imagine a pricing model that not only considers capacity, but also takes social relationships into accounts, e.g. whether or not a popular person or a friend of yours purchased the offer.

It is evident that the long road of our journey still lies ahead.

5 Conclusions

The following core features characterize our approach:

- 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. The alignment of content and channel is achieved through a weaver that aligns ontological items with channels.
- These ontologies are not case-specific, but model a certain vertical domain such as research projects, associations, accommodations, restaurants, bars, touristic events and services, etc. Therefore, these ontologies and their channel alignments can be reused on a larger 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, and semantic based channels, using various formats and vocabularies.

Based on our approach, ESWC, PlanetData, and STI International are now managing their on-line appearance. Around 300 concepts and properties, 500 channels, i.e., around 150,000 potential content-to-channel mappings are run by a

very small dissemination team. Currently, we are performing additional 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 eTourim, where millions of hotels are desperately waiting for a scalable dissemination strategy, given the fact that soon, around 50% of all room bookings will be done on-line.

Acknowledgements. We would like to thank Anja Bunnefeld, Johannes Breitfuss, Carmen Brenner, Alice Carpenter, Anna Fensel, Michael Fried, Mark Greaves, Marko Grobelnik, Martin Hepp, Lyndon Nixon, Ina O' Murchu, Alexander Oberhauser, Simeona Pellkvist, Elena Simperl, Corneliu Valentin Stanciu, Ioan Toma, and Alexander Wahler for contributions to early drafts of this paper.

6 References

[Amersdorffer et al. (eds.), 2010] D. Amersdorffer, F. Bauhuber, R. Egger, and J. Oellrich (eds.): *Social Web im Tourismus*, Springer, 2010.

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

[Bauereiß et al., 2011] T. Bauereiß, B. Leiter, and D. Fensel: Effective and Efficient On-lin Communication, STI TECHNICAL REPORT 2011-12-01, 2011. http://www.sti-innsbruck.at/TR/OnlineCommunication

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

[Bizer & Schultz, 2010] C. Bizer and A. Schultz: The R2R Framework: Publishing and Discovering Mappings on the Web. In Proceedings of the 1st International Workshop on Consuming Linked Data (COLD 2010), Shanghai, November 2010. http://www4.wiwiss.fu-berlin.de/bizer/r2r/

[Bonchi et al., 2011] F. Bonchi, C. Castillo, A. Gionis, and A. Jaimes: Social Network Analysis and Mining for Business Applications, *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3), 2011.

[Corlosquet et al, 2009] 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.

[Farris et al., 2010] 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.

[Fensel, 2001] D. Fensel: Ontologies: Dynamic Networks of Formally Represented Meaning. 2001. http://sw-portal.deri.at/papers/publications/network.pdf

[Finzen et al., 2010] J. Finzen, H. Kasper, and M. Kintz: Innovation Mining, Frauenhofer Verlag, 2010.

[Franklin et al., 2011] M. J. Franklin, D. Kossmann, T. Kraska, S. Ramesh, and R. Xin: CrowdDB: Answering queries with crowdsourcing, *SIGMOD '11 Proceedings of the 2011 international conference on Management of data*, 2011.

[Garcia et al. 2008] 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.

[Helbing & Konitzer, 2010] T. Helbing and M. Konitzer: Das Ohr am Puls des Internets. In [Amersdorffer et al. (eds.), 2010].

[IDEAS, 2005] THE BASICS OF REVENUE MANAGEMENT, Integrated Decisions and Systems, Inc., 2005. http://www.adhp.org/pdf/1-theBasicsofRM.pdf

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

[Kasper et al., 2010] H. Kasper, M. Dausinger, H. Kett, and T. Renner: Marktstudie Social Media Monitoring Tools, Frauenhofer Verlag, 2010.

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

[Kimes, 2000] 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,

[Kobilarov et al., 2009] 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.

[Kosala & Blockeel, 2000]. R. Kosala. and H. Blockeel: Web mining research: A survey, *ACM SIGKDD Explorations Newsletter*, 2(1):1-15, 2000.

[Lermann et al., 2011] K. Lerman, S. Intagorn, J.-H. Kang, R. Ghosh: Using Proximity to Predict Activity in Social Networks. CoRR, Vol. abs/1112.2755. http://arxiv.org/abs/1112.2755

[Lermann et al., 2011] K. Lerman, S. Intagorn, J.-H. Kang, R. Ghosh: Using Proximity to Predict Activity in Social Networks. CoRR, Vol. abs/1112.2755. http://arxiv.org/abs/1112.2755

[Li & Bernhoff, 2008] C. Li and J. Bernhoff: *Groundswell: winning in a world transformed by social technologies*, Forester research, Harvard Business School Publishing, 2008.

[Ma et al., 2006] 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.

[Morales-del-Castillo et al., 2009] 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.

[Mulpuru et al., 2011] S. Mulpuru, H. H. Harteveldt, and D. Roberge: Five Retail eCommerce Trends To Watch In 2011, Forrester Research Report, January 31, 2011.

[Schultz et al., 2011] A. Schultz, A. Matteini, R. Isele, C. Bizar, and C. Becker: LDIF – Linked Data Integration Framework. In Proceedings of the 2nd International Workshop on Consuming Linked Data, Bonn, Germany, October 2011. http://www4.wiwiss.fu-berlin.de/bizer/ldif/

[Sheth et al., 2002] 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.

[Solis, 2012] B. Solis: *The Rise of Digital Influence*, Altimeter Group, 2012.

[Staab & Studer (eds.), 2009] S. Staab and R. Studer (eds.), Handbook on Ontologies, Springer Verlag, Heidelberg. 2nd edition, 2009.

[Stavrakantonakis et al., submitted] I. Stavrakantonakis, A.-E. Gagiu, H. Kasper, I. Toma, and A. Thalhammer: An approach for evaluation of social media monitoring tools, submitted.

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

[Wang et al., 2011] G. Wang, C. Wilson, X. Zhao, Y. Zhu, M. Mohanlal, H. Zheng and B. Y. Zhao: Serf and Turf: Crowdturfing for Fun and Profit CoRR, Vol. abs/1111.5654, 2011. http://arxiv.org/abs/1111.5654