Enabling Scalable Multi-Channel Communication through Semantic Technologies

Ioan Toma, Dieter Fensel, Andreea-Elena Gagiu, Ioannis Stavrakantonakis, Anna Fensel, Birgit Leiter, Andreas Thalhammer, Iker Larizgoitia and Jose Maria Garcia
STI Innsbruck, University of Innsbruck
ICT Technologiepark, Technikerstrasse 21a, 6020 Innsbruck, Austria
e-mail: {firstname.lastname}@sti2.at

***Abstract* — With the advance of the Web in the direction Social Media the number of communication possibilities has exponentially increased bringing new challenges and opportunities for companies to build and shape their reputation online as well as to engage and maintain the relationships to their customers. In this paper we describe how semantic technologies enable scalable, effective and efficient on-line communication. We illustrate four different ways in which semantics can be used for this purpose. First, we discuss semantic analysis of communication items based on 'classical' semantic, such as natural language processing. Second, we look at semantics as a channel, viewing Linked Open Data vocabularies not only as terminological assets but as communication channels. Third, semantics provide the methodologies and tools for content modeling by means of ontologies. Finally, semantics through semantic matchmaking enable semi-automatic assignment and distribution of content to channels and vice-versa.**

Keywords – multi-channel communication, semantics, semantic technologies, social media

I. INTRODUCTION

A decade ago companies embraced the Web as a novel means to advertise product and service offerings via information they published on their Web sites. Today, on the user-empowered (mobile) Web, the focus has shifted from the individual, self-controlled Web sites towards Social Media platforms, including forums, (micro)blogs, social networks, and increasingly mobile apps, affecting the ways in which companies can build and shape their own reputation, as well as the relationships to their customers and the general masses.[1]

In this exceedingly dynamic and decentralized environment, essentially anyone can express their opinions about organizations and their offerings, potentially influencing the perception by the online public and the purchase behaviour of a significant share of the market. Nowadays it has become much easier, and at the same time mandatory, to listen and engage in dialog with customers, to promote brands and offerings, and to conduct business transactions through an ever-increasing number of communication channels. The ever growing number of communication channels brought in by Social Media the problem becomes 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 become more and more granular* and increasingly dependent on the identity of the receiver and the context of the communication.

The constantly growing opportunities in communication provided by Social Media can be both an enabler and a burden. Mulpuru [23] calls the later aspect “the growth of the multichannel monster”. Being present in a multitude of different channels requires the effective management of a very large number of adapted contents, formats, and interaction patterns fulfilling the communication and cooperation needs of distributed target groups. Organizations require new skills and more efficient access means to scale and filter the exponentially increased offer.

This paper discusses how semantic technologies can be used to address the challenges introduced by exponential growth of communication channels. More precisely we use semantic technologies: (1) to process and analyze the communication content, to identify the sentiment and changes in sentiment, to extract opinions be means of Natural Language Processing, (2) to utilize Semantic and Linked Data vocabularies as additional channels for communication, (3) to model communication content by means of ontologies and (4) to enable semi-automatic matchmaking of content to channels and vice-versa.

The remainder of the paper is organized as follows. Section II describes in details the four ways semantics technologies are used for scalable multi-channel communication. Section III explains how the four ways of using semantic technologies are integrated into our proposed approach for multi-channel communication, describes the use case used to showcase our solution and provides details about the implementation work. Section IV discusses the related work. Finally, Section V sketches the future work and concludes the paper.

[1] [1] ‘Business gets social’ by Carol Rozwell, Gartner, 2011

II. Four Roles For Semantic Technologies

Semantic Technologies [5] 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. This section discusses which Semantic Technologies to use and how to use them to achieve efficient and effective multi-channels communication.

A. Semantic Analysis

Semantic Analysis enables computers to "understand" the natural language statements in a communication act. Semantic Analysis is commonly implemented by using Natural Language Processing[2] (NLP) techniques. In broad terms, NLP aims to design and build software that analyses, understands and generated languages that humans use naturally, in order to eventually enable users to address their computers as though they were addressing another person[3]. A variety of NLP tasks, including Topic detection, Named Entity Recognition, Sentiment Detection, Opinion Mining, etc. are very relevant for analyzing the content of communication and taking automatic decisions. NLP based semantic analysis enables discovery of facts in texts, recognition of mentioned entities, enriching of factual information with background knowledge, identifying the sentiment convey in the content, changes in sentiment, extraction of opinions, etc. Among these very important are opinion mining and sentiment analysis. Opinions and sentiments are determined using elements of computational linguistics, text analytics and machine learning (e.g. latent semantic analysis, support vector machines, etc.). The sentiment score may be extremely important in evaluating the enterprise's brand in a large data set on user generated mentions, as well as enable the enterprise to filter the content based on positive and negative comments. This process additionally enables the enterprise to single out the themes and issues that have determined the development of the discovered sentiment. Although automated sentiment technology cannot reach the quality of a human annotator, it offers advantages such as speed (which is near real-time) and the ability to process a large quantity of data. In addition, the techniques are tireless, fast, and consistent which can be improved over time [19].

B. Semantic Channels

Channels are first class components of any communication model (for an early communication model see [25]). They provide the medium through which the communication content is transmitted. The recent paradigm shift with respect to Semantic Technologies towards the Web of Data vision and its implementation Linked Open Data (LOD)[4] provides new opportunities to integrate more data centric communication channels. The Linked Open Data initiative focuses on publishing datasets using the Resource Description Framework (RDF) [18], the metadata model primarily used on the Semantic Web, and creating links between the entities in these datasets. Central to the Linked Open Data initiative are vocabularies and languages.

Vocabularies on the other hand reflect the need of having terms to describe the data and the content that is disseminated and shared in the Web. The aim of adding annotations to the data and the content is to give meaning and semantics to them in order to leverage the content to self-explanatory content which can be understood not only by humans but also by machines. Thus, in addition to predefined formats, we need to reuse predefined *vocabularies* to describe our data to enable semantic-based retrieval of information.[5] Widely used vocabularies are:

- *Dublin Core* "… set of metadata elements provides a small and fundamental group of text elements through which most resources can be described and catalogued."[6]
- "*The Friend of a Friend (FOAF)* project is creating a Web of machine-readable pages describing people, the links between them and the things they create and do"[7]
- *GoodRelations* is a vocabulary for publishing details of products and services optimized towards search engines, mobile applications, and browser extensions.[8]
- *Schema.org* "… provides a collection of schemas, i.e., html tags that webmasters can use to markup their pages in ways recognized by major search providers."[9,10]

We interpret these vocabularies as channels. If we map information item in such a vocabulary, it can be understood by other agents that are common with this vocabulary. 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. Content can be exported or imported to or from it. *Therefore, LOD vocabularies are means to disseminate and share information and not means to model information.*

Formats refer to the languages that are needed in order to use the vocabularies to annotate content and integrate the

[2] http://en.wikipedia.org/wiki/Natural_language_processing
[3] http://research.microsoft.com/en-us/groups/nlp/

[4] http://linkeddata.org
[5] More than a hundred of them are listed at http://labs.mondeca.com/dataset/lov/index.html.
[6] http://en.wikipedia.org/wiki/Dublin_Core
[7] http://www.foaf-project.org/
[8] http://www.heppnetz.de/projects/goodrelations/
[9] http://www.schema.org/ and http://schema.rdfs.org/
[10] Further LOD vocabularies are Event Ontology, an Organization ontology, AIISO (Academic Institution Internal Structure Ontology), DOAP (Description of a Project), Project Documents Ontology, SIOC (Semantically-Interlinked Online Communities) Core Ontology, SWC (Semantic Web Conference ontology), SWRC (Semantic Web for Research Communities), and VIVO.

semantics into the data. Thus, there are formats that are used to describe the schemas of the vocabularies on which the vocabularies are built. Moreover, there are formats for the data inclusion in documents, which are the actual means of enriching the content with terms extracted from various vocabularies. There are various formats available for data inclusion in documents. Finally, there are formats that can be used to make data from repositories directly available in the Web.

To satisfy the needs of different application fields, W3C offers a large variety of techniques to describe and define different forms of vocabularies in a standard format, which include RDF [18] and RDF Schemas [1], Web Ontology Language (OWL) [27], and the Rule Interchange Format (RIF) [17]. Choosing amongst these technologies is largely dependent on the rigor required by the specific application. Besides the formats for schemas mentioned before, additional formats are needed for inclusion of data into documents and for direct data publication. Embedding formal metadata in documents can be done using mark-up formats, such as RDFa, Microformats, or Microdata. For direct data publication and making data from repositories in the Web formats such as RDF and SPARQL [24] are used. Moreover Linked Data principles defined by Tim Berners-Lee [1] should be followed, namely: (1) use URIs as names for things; (2) use HTTP URIs to enable users to look up the names for things; (3) provide useful information for looking up a URI (using standards, such as RDF and SPARQL) and (4) include links to other URIs to enable the discovery of more things.

C. *Semantic Content Modeling*

Semi-automation of online communication processes is only possible if content can be understood not only by human agents but by machines as well. Semantic Technologies in general and Ontologies in particular provide the means to conceptualize and share content, a prerequisite for automation. Gruber [11] defines and ontology as a "formal, explicit specification of a shared conceptualization". An ontology is used by an agent, application or other information resource to declare what term the agent uses and what the terms mean [26]. Making this information available enables the possibility for high fidelity semantic communication – agents can communicate, share meaning with other agents and understand the meaning of applications, databases and other information resources.

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 employed 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." [7]. We achieve this by mapping domain content with LOD vocabularies when we see a gain in broadening our range of communication through them.[11] The content itself is often better modeled in a well-defined and carefully crafted Ontology targeted for the specific domain and range of tasks it should support. Content modeling only has to be done once using Domain Ontology that is understandable by the domain experts. In this way content becomes reusable and can be presented in various ways. Last but not least content and channels should be brought together via an interweaving process (see Section III, C).

D. *Semantic Matchmaking*

Using a semantic matchmaking channels and content are matched automatically. Using semantics the semi-automatic generation of weaving content and channels logic becomes possible. Semantic matchmaking enables automatic distribution of semantically described content into the right channels as well as collection of feedback and statistics from the channels (See Figure 1).

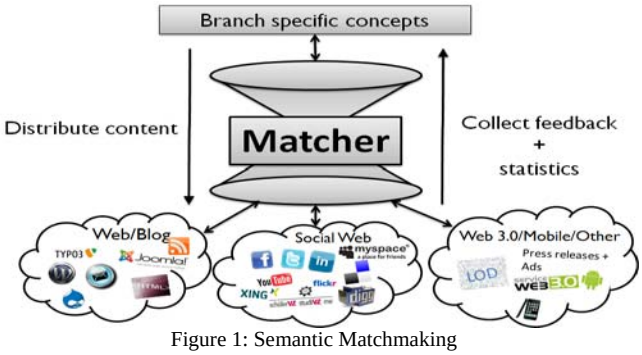

Figure 1: Semantic Matchmaking

Based on its semantic annotations, content is matched, transform and updated to the right channels. Typical transformations include shortening of text to fit into tweets, attachment and resizing of pictures, transformation and adaptation of videos or slides where needed.

III. A SEMANTIC BASED MULTI-CHANNEL COMMUNICATION APPROACH

The core idea of our approach for scalable, cost-sensitive and effective online multi-channel dissemination is to introduce a layer on top of the various Internet based communication channels that is domain specific and *not* channel specific[12]. The domain specific ontologies that provide the information models are one of the core elements of our approach. Additionally, four other core elements are

[11] "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.

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

required to build a scalable multi-channel communication solution, namely:

- 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;
- a library of implemented wrappers for actual channel instances; and finally
- a set of communication patterns, i.e. reusable templates for management of communication.

What is essential is to *distinguish* the communication or channel model from the conceptual descriptions of the information, in analogy to cascade style sheets (CSS) that separate the content from its presentation. Our approach requires the creation of a communication model (i.e., an increasingly complete model of channels), and knowledge models for each vertical (for the tourism domain this includes hotels, restaurants, tourist events, etc.), and finally linking the knowledge model with the communication model through a weaver that weaves concepts with channels. Data and information can be expressed at the conceptual level, which the domain expert understands. The knowledge models are formalized using ontologies [6], [7] and are not full descriptions of the verticals but rather are focused on the information chunks that are disseminated about it. In this section, we describe how the content and the multitude of communication channels are being managed and interweaved as part of our semantic based multi-channel communication solution.

A. *Managing the Content*

In our approach, information items being communicated are specified using a domain terminology that is easily understood by domain experts. Such a domain specific terminology is not an exhaustive formalization of the domain, but rather includes the concepts that are most frequently involved in the various acts of communication. The domain specific terminology is formalized as an ontology. For the touristic domain, one of the core domain in which we apply our solution, we have developed the Accommodation Ontology [13] [13]. The Accommodation Ontology is an extension of GoodRelations[14], which was extended with additional vocabulary elements for

- describing hotel rooms, hotels, camping sites, and other forms of accommodations, their features, and
- modeling compound prices as frequently found in the tourism sector, e.g. weekly cleaning fees or extra charges for electricity in vacation homes based on metered usages.

GoodRelations is a very popular e-Commerce ontology for annotating offerings and other aspects of e-commerce in the Web and is the only OWL DL ontology officially supported by both Google and Yahoo. Only for very few features, readily standardized conceptual elements are provided by the Accommodation Ontology ontology, like: acco:occupancy, acco:occupancyAdults, acco:petsAllowed, acco:occupancyMinors, and acco:size.

We are using semantic technologies to support the overall management of content dissemination in a multi-channel and bi-directional communication setting. We use vertical domain models, which are shared and reused in a vertical area instead of being used for a single application only. The vertical domain models are formalized as ontologies in RDFS [1] or OWL [27]. As mentioned in Section II, Linked Open Data[15], RDF [18] vocabularies are not seen as models in our approach but rather as channels for dissemination. Popular vocabularies, broadly used by organisations in the market and research institutes, such as *GoodRelations* and *Schema.org*, can serve for this purpose.

B. *Managing the Communication Channels*

A core feature of our approach is to abstract information from the underlying online communication channels. The channels are the vast amount of on-line communication possibilities. In our view, channels are means of exchanging information in the on-line space, and include a wide range of online communication possibilities, such as static dissemination, dynamic dissemination, dissemination through sharing, dissemination through collaboration, dissemination through group collaboration and semantic dissemination (see [8] section 2.2 for extended details).

Static dissemination refers to the unchangeable (or almost unchangeable) aspect of the information dissemination. Information broadcasted statistically rarely changed, or changes at very large time intervals. Examples of static challenges include *websites* (homepages), *Wiki* pages, printed press, and *Content Management Systems* (CMS). In addition to static dissemination channels several tools exist that to measure uptake and impact of the information dissemination. *Google Analytics* [16] and *Yahoo! Web Analytics*[17] are web analytics solution providing insight into website traffic and marketing effectiveness.

Dynamic dissemination refers to dissemination approaches in which the information is available or relevant only for a limited period of time. Most Web2.0 channels fall into this category including blogs, news articles, microblogs, etc. Very relevant dynamic dissemination channels are news feeds such as *RSS feeds* and *microblogs*. Particularly, an important channel in this dimension is Twitter, whether it is used to read or receive news to building virtual communities and opening online conversations. Last but not least, dynamic dissemination channels include *email*, one of the most reliable medium for e-Commerce driver and customer relationship builder, *blogs*, popular medium to share ideas

[13] http://ontologies.sti-innsbruck.at/acco/ns.html
[14] http://www.heppnetz.de/projects/goodrelations/
[15] http://linkeddata.org/
[16] http://www.google.com/analytics/
[17] http://web.analytics.yahoo.com/

and *chatting and instant messaging* for instantly communicating and disseminating information.

Dissemination and communication through sharing includes some of the most popular Web2.0 channels. Examples for sharing are YouTube, Vimeo for sharing videos, Flickr, Pinterest for sharing pictures, SlideShare for slides, Delicious, Digg for social bookmarking, etc.

Dissemination through collaboration channels include tools such as *wiki*, some of the most popular and important tools for collaborative development of information and shared knowledge. Dissemination through group collaboration includes channels that support sharing and exchanging information but also to collect feedback or discuss certain issues. Social networks such as Facebook, Google+, LinkedIn and Xing are well-known example of channels from this category. *Facebook*, with its more that 955 million monthly active users, is especially notable in the realm of social networks[18] offering many possibilities through which to disseminate information. Social networks such as *XING* and *LinkedIn* may be very useful for finding new employees or business partners.

Finally, semantic dissemination can be realized through the multitude of semantic channels described in Section II, B.

The channels mentioned above are very heterogeneous in nature, having various interfaces, allowing different content type, and supporting different interaction modes- just to name a few of the heterogeneity of aspects. Our solution provides an abstraction from these channels, integrating and personalizing them. Basic interaction with each of these channels, e.g. the ability to read from the channel and write to the channel, is also supported. Finally, our solution supports the aggregation of channels into composed communication structures.

C. Weaving the Content and Communication Channels

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 modelling 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 the following elements:

1. ***An information item***: 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.

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 intentional 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.
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.

Based on these features, the publication process can be formally defined in a way in which the multi-channel publication is managed automatically. Figure **2** shows a sample definition (not using any formalism, simply for educational purposes) of two tuple sets describing the weaving process for the publication of a hypothetical Event in two different channels (Twitter and Flickr), these tuples are processed by the weaver and the actions are launched accordingly. We can see the flexibility of this approach in many ways. For example, in the first one, the editor is a human (the hotel animator who has to define the event and follow the pre-established internal protocol), while on the executor side we have references to software components that will publish the content in the appropriate channel. The definition of processing rules can also be interesting, as defined, for example, in the Twitter channel. In this example, the description of the event should be less than 140 characters, and a summarization is usually needed (adaptation). In the case of summarization, the rule has a different purpose; it serves to check that the event has an image defined, because it is mandatory to publish it on Flickr. Scheduling options are also defined to enable control of the timing of publications and additional information as meta-information. Reflecting the state of the tuple serves as a control mechanism for the weaving process.

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

[18] Information taken from: http://newsroom.fb.com/content/default.aspx?NewsAreaId=22

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 a major contribution to the scalability of our multi-channel communication approach.

Information Item	Event
Content Editor	Hotel animator
Editor's Interaction Protocol	Event Organization Protocol
Information Type	Instance
Processing Rule	if event.description > 140, summarize
Channel	twitter
Scheduling Information	once a week until event.date
Executor	TwitterChannelConnector
Executor's interaction protocol	TwitterChannelConnector API
Meta-information	active

Information Item	Event
Content Editor	Hotel animator
Editor's Interaction Protocol	Event Organization Protocol
Information Type	Instance
Processing Rule	if event.picture == null, deactivate
Channel	flickr
Scheduling Information	once
Executor	FlickrChannelConnector
Executor's interaction protocol	FlickrChannelConnector API
Meta-information	active

Figure 2: Tuple set example of the weaving process (no formalisms applied)

D. Communication Patterns

To further facilitate communication, collaboration and value exchange, we propose a set of on-line communication patterns that can be reused to facilitate the workflow of the communication. Communication patterns are metaphors adopted from the software engineering domain (i.e. design pattern). The communication patterns paradigm stems from the need of formalising the communication that is taking place between the various stakeholders in different domains. For example in the tourism domain, hoteliers disseminate a huge variety of content to the customers and potential guests in order to trigger their attention and engage with them. Thus, hoteliers should be able to decide before any campaign and package offering dissemination, which is the most effective way to follow and which communication channels should be used. Moreover, the content and the media attached to the messages vary according to the goal of the communication. At a later stage, the sender of the message should react and interact with the replies to the initial message that were sent by the users. Therefore, the aforementioned workflow of communication could be facilitated by providing to the businesses that are managing campaigns and disseminating news as part of their marketing efforts, like hoteliers, with means to structure and model the communication processes. Predesigned communication templates that address specific needs and issues could be reused by the people in order to make easier and better decisions than they used to make in the past. This idea is the main objective of the communication patterns initiative.

The communication patterns can be understood as an implementation of business processes. A business process is a set of independent activities that need to be performed in response to a business event, to achieve a business objective [10]. The communication patterns paradigm tries to recognise patterns in the communication between the agents and propose solutions and workflows that could be followed in order to have an effective cooperation and a positive impact on both sides. In this respect, the communication patterns could be considered as a set of predefined templates of business processes that address certain needs of business entities (e.g. hoteliers) in a specific domain (e.g. tourism) regarding the communication with customers and potential customers.

In terms of representation of communication patterns, two aspects need to be considered. The first one is the definition of the communication patterns as workflows of activities by the domain experts. The second part of workflows descriptions is related to the definition of the workflows in a way that is appropriate for the workflow engine. UML can be used to model the requirements for a communication process and a communication pattern from the domain expert to the workflow engine administrator project via the usage of UML activity diagrams.

Figure 3 depicts an example of package offer dissemination by using UML.

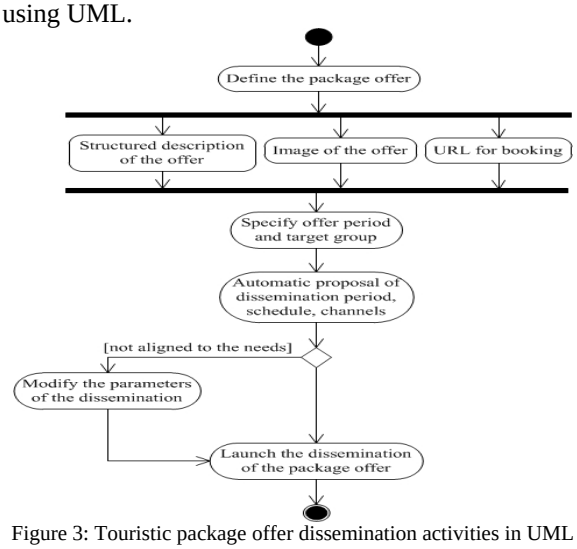

Figure 3: Touristic package offer dissemination activities in UML

For the definition of the workflows in a way that is appropriate for the workflow engine, workflow specification descriptions in languages such as YAWL [19] should be

[19] http://www.yawlfoundation.org/pages/resources/patterns.html

generated from the previous created UML diagram in order to feed the workflow engine with the related tasks.

E. Use Case and Implementation

We are applying our solution for multi-channel communication in the e-Tourism domain. The challenges and opportunities that lie in this field are the reasons that we are firstly targeting to support the tourism domain with the application of approach. The internet and web-based communication and booking channels are becoming increasingly important in today's completive tourism domain. More than 55% of all tourists in Central Europe inform themselves on-line about a certain destination before booking, and more than 27% of all tourists in this area use internet-based booking channels for reserving their tourism plans[20]. Currently, the Web is one of the most important sources to plan trips, holidays and business travel. The main objective of our work is to help the hotelier in dealing with the challenge to improve and maintaining his communication needs in a world with an exploding number of channels in order to maintain or better increase his market share (i.e., the number of bookings and the attached price) by keeping the related transactions costs for on-line communication and booking manageable.

The implementation of our approach is under development as part of the Seekda Social Agent (SESA) project[21] with the ultimate goal of enabling communication, collaboration and value exchange (i.e. booking) at scale for small and medium enterprises active in the tourism domain. In the scope of SESA, we specify and realize the concepts of channel model [8], weaving process of content and channels and communication patterns. SESA is delivering a platform that will help businesses in the tourism domain in dealing with the challenge of improving and maintaining their communication needs in order to engage with their customers in a proper and timely manner, and to effectively and efficiently support value exchange. SESA integrates various social media channels including Facebook, Twitter, LinkedIn, YouTube and Flickr, disseminates information items with one click through a multitude of channels, shows feedback collected from customers, and supports engagement and value exchange through a multitude of on-line interaction possibilities based on the use of *semantic technology*.

IV. Related work

We see two specifically related areas: *Ontology-based content management systems (CMSs) for websites* and *Semantic matchmaking of senders and receivers of content.*

The field of *semantics-based or enhanced CMSs* has already been quite thoroughly explored. As an early pioneer for ontology-based website management, OntoWebber system [14] introduces an integration layer that adapts to different data sources. This is related to our weaver concept but, in contrast, the weaver adapts to different channels rather than to different information sources. [9] introduces "The Rhizomer Semantic Content Management System" which integrates services with metadata browsing, editing, and uploading, continuing their earlier work on the Knowledge Web portal. [4] proposes a Linked Data extension for Drupal that enables content annotation with RDFa and provides a SPARQL endpoint. As reported in [2], BBC's World Cup 2010 site[22] 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. Compare to these approaches, our approach supports 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. Obviously we can only select a small sample of approaches in this area, which focus on matchmaking in regard to content. [15] 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. The system introduced in [20] 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 [22] 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. 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. Fortunately, a large number of such web analytical toolkits already exist, [16] 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.

V. Conclusion and Future Work

In this paper, we discussed how semantic technologies enable scalable, effective and efficient on-line

[20] http://www.newmediatrendwatch.com/regional-overview/103-europe?showall=1

[21] http://sesa-project.sti2.at/

[22] http://www.bbc.co.uk/worldcup

communication. We illustrate four different ways in which semantics can be used for this purpose. More precisely we use semantic technologies: (1) to process and analyze the communication content, to identify the sentiment and changes in sentiment, to extract opinions be means of Natural Language Processing, (2) to utilize Semantic and Linked Data vocabularies as additional channels for communication, (3) to model communication content by means of ontologies and (4) to enable semi-automatic matchmaking of content to channels and vice-versa. We also proposed an approach that combines the four ways of using semantic technologies into a unified approach characterized by the following core features: (i) we use ontologies to model content in order to have a representation layer independent from the communication channel. The alignment of content and channel is achieved through a weaver that aligns ontological items with channels; (ii) ontologies are not case-specific, but model a certain vertical domain; (iii) our approach is bi-directional; (iv) we support in an integrated fashion, the dissemination via traditional web channels, Web 2.0, and semantic based channels, using various formats and vocabularies.

For our approach, semantics is a corner stone but requires many additional services and layers to actually provide its potential. Together with seekda[23] we are currently focusing on the eTourism domain, however, other verticals may follow. In general, we target domains (verticals) with many SMEs that need to intensively interact with their customers on-line.

ACKNOWLEDGMENT

We would like to thank all the members of the Online Communication (http://oc.sti2.at/) working group and the SESA Project (http://sesa-project.sti2.at/) for their valuable feedback and suggestions.

REFERENCES

[1] T. Berners-Lee, (18.06.2009). Linked Data. Retrieved from W3.org: http://www.w3.org/DesignIssues/LinkedData.html. Last accessed 17.05.2012.

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

[3] D. Brickley, R. Guha, and B. McBride, "RDF Vocabulary Description Language 1.0: RDF Schema", W3C Recommendation. Retrieved from W3C Recommendation: http://www.w3.org/TR/rdf-schema/, 2004.

[4] 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.

[5] J. Domingue, D. Fensel, and J. A. Hendler. (2011). Handbook of Semantic Web Technologies (Vol. 1). Springer.

[6] D. Fensel. Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce, Springer-Verlag, Berlin, 2001. 2nd Edition, Springer-Verlag, Berlin, 2003.

[7] D. Fensel. Ontologies: Dynamic networks of formally represented meaning, Vrije University: Amsterdam, 2001.

[23] http://seekda.com/

[8] A. Fensel, D. Fensel, A.-E. Gagiu, J. Kaiser, I. Larizgoitia, B. Leiter, I. Stavrakantonakis, A. Thalhammer, and I. Toma, I, "How to Domesticate the Multi-Channel Communication Moster", Technical Report available at: http://oc.sti2.at/results/white-papers/how-domesticate-multi-channel-monster, 2012.

[9] A. Fensel, D. Fensel, B. Leiter, and A. Thalhammer. (2012). Effective and Efficient Online Communication: The Channel Model, In Proceedings of International Conference on Data Technologies and Applications, SciTePress, Rome, Italy, 25-27 July, 2012.

[10] 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.

[11] T. Gruber. A translation approach to portable ontology specifications. Knowledge Acquisition 5:199-220, 1993.

[12] J. Han, and J. Mills. Zero acquaintance benchmarking at travel destination websites: what is the first impression that national tourism organizations try to make? International Journal of Tourism Research, 405-430, 2006.

[13] M. Hepp, "GoodRelations Language Reference" 2011, retrieved from GoodRelations: http://www.heppnetz.de/ontologies/goodrelations/v1 (last accessed 28.04.2013).

[14] 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.

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

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

[17] M. Kifer. Rule Interchange format: The framework. Journal of Web reasoning and rule systems, Springer, 1-11, 2008

[18] G. Klyne, J. Carroll, and B. McBride, "Resource Description Framework (RDF): Concepts and Abstract Syntax", W3C Recommendation. Retrieved from W3C Recommendation: http://www.w3.org/TR/rdf-concepts/, 2004.

[19] J. Kmetz.: Measuring Social Sentiment: Assessing and Scoring Opinion in Social Media. Retrieved from VisibleTechnologies.com: http://www.visibletechnologies.com/resources/white-papers/measuring-sentiment/.

[20] 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.

[21] B. Martens and G. Turlea. The drivers and impediments for online cross-border trade in goods in the EU. Digital Economy Working Paper 2012/1.

[22] 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.

[23] S. Mulpuru, H. Harteveldt, and D. Roberge, "Five retail eCommerce trends to watch", in 2011, Forrester Research Report, January 31, 2011.

[24] E. Prud'hommeaux, and A. Seaborne. SPARQL Query Language for RDF. Retrieved from W3.org: http://www.w3.org/TR/2006/WD-rdf-sparql-query-20061004/ (last accessed 17.05.2013).

[25] C. Shannon, and W. Weaver. The mathematical theory of communication. . Urbana, IL: University of Illinois Press, 1949.

[26] M. Uschold, and M. Gruninger. Ontology and semantics for seamless connectivity. SIGMOD Record 33(3), 2004.

[27] F. Van Harmelen, and D. L. McGuinness. OWL Web Ontology Language (OWL): Overview. W3C Recommendation, retrieved from W3C Recommendation: http://www.w3.org/TR/owl-features/.