Social Media Integration into the GS1 Framework

Irena Pletikosa Cvijikj, Florian Michahelles, Elgar Fleisch

Auto-ID Labs White Paper WP-BIZAPP-058

January 2011

Irena Pletikosa Cvijikj
Researcher and PhD Student
ETH Zürich

Dr. Florian Michahelles
Associate Director of the
Auto-ID Labs Zurich/St. Gallen
Manager of the labs of Prof.
Fleisch at D-MTEC
ETH Zürich

Prof. Dr. Elgar Fleisch
Professor of Information and
Technology Management at
ETH Zürich and
University of St. Gallen (HSG)

Contact:
Irena Pletikosa Cvijikj
ETH Zürich
Phone:+41 44 632 86 24
Fax: +41 44 632 17 40
ipletikosa@ethz.ch
www.im.ethz.ch
Abstract

Social media has changed the way content is generated on the web. Rather than being just passive consumers, users became active participants by sharing information, experiences and opinions with each other. In parallel, with the continuous growth of the number of smartphone users, brand owners are being faced with the challenge of integration of social media and smartphone technologies to effectively connect and engage with their consumers. To contribute in this direction we propose a solution for integration of social media into the GS1 framework based on the existing architecture of the GS1 MobileCom/B2C solution. In addition, we provide detailed architecture of the components and tools involved in the proposed ‘Social Media Content Provider’ which should be integrated as a new building block into the GS1 architecture. Finally, we evaluate the validity of the solution by discussing the opportunities emerging from the presented approach for both, brand owners and consumers.

1. The Rise of Social Media

At the end of the 2011, one might summarize the most important events by looking at the buzzwords that have marked the year. According to an article by David Ball [4], the first place on the list belongs to the Social Media Marketing, denoting a new era of B2C communication where the control moves from companies to consumers. The list continues with the concept of Social Proof, i.e. placing a value on the friends’ opinions, followed by Microblogging which could be described as broadcasting one’s life online, then Viral, the new way of describing the well established concept of online word-of-mouth (WOM), and Social Media Optimization, i.e. promoting company’s online presence on social media by providing possibility to share its content. In summary, it is all about the Social Media.

So what exactly is social media? In academic literature social media is defined as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content.” [32] This definition is based upon two additional concepts, the Web 2.0 referring to the underlying technological platforms and the User Generated Content (UGC) which, in a broad sense, can be seen as “the sum of all ways in which people make use of social media.” [32] In other words, rather than being just passive consumers, users are becoming active participants by sharing information, experiences and opinions with each other [7].

While the concept of social media might be dated as late as 1979, when Tom Truscott and Jim Ellis from Duke University created the Usenet, an online discussion system that allowed posting of public messages [32], the social media as we know it today probably emerged
about 20 years later, with the creation of the first social network (SN) SixDegrees.com [11]. A SN can be defined as “web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system.” [9] Following the period of several smaller launches, the significant expansion of SNs started in 2003 with LinkedIn, MySpace, Flickr, etc., many of which are still in use today.

The major leap happened with Facebook in early 2004 [11]. Facebook differed from previous SNs by preventing public access to the user profile pages. Instead, a friend request confirmation was needed to grant reciprocal access to personal data. At the time of writing, Facebook is the largest SN with more than 800 million active users [23]. With such large numbers of users and activities, and a documented loss of consumer trust in traditional advertising [13], it is no wonder that SNs have attracted the attention of advertisers, brand owners and retailers, with predictions of 3.93 billion USD spent for advertising on SNs in 2012 [19], thus making the Social Media Marketing the buzzword of the year.

2. Social Media Marketing

Social media marketing, a form of WOM marketing, but also known as viral marketing, buzz, and guerrilla marketing, is intentional influencing of consumer-to-consumer communication through professional marketing techniques. This is not to be seen as a replacement for the traditional marketing techniques but rather as an additional marketing channel that could be integrated with the traditional ones as a part of the marketing mix. The advantage of this new electronic channel is that it can be used to communicate globally and enrich marketing toward consumers, at the personal level [10]. Through users’ feedback or by observing conversations on social media, a company can learn about customers’ needs, potentially leading to involvement of members of the community in the co-creation of value through the generation of ideas [38].

The rise of social media has democratized the traditional marketing communication leading to fundamentally altered marketing’s ecosystem of influence [47]. In such ecosystem, consumers started dictating the nature, reach and context of marketing messages, extending their effect through the shared content [29]. They have started engaging in brand related communication with or without the company’s involvement. In response, companies have evolved their customer approach, shifting from traditional one-to-many communication to a one-to-one approach and offering contact or assistance at any time through one or more platforms such as Facebook, Twitter, MySpace, etc. [25]. This approach has shown that blogs can be useful for generation of sales leads, smartphones provide possibility for contextual and location-based interactions and company’s videos on YouTube drive the sales [14].
Still, viral marketing on social media has not yet reached the high expectations set [13]. Today, on account of the newness, companies experiment with many different forms of interaction, sometimes with great success, e.g. Nutella found a communication tone that helped it become one of the most successful brands on Facebook. By contrast, poor understanding of the medium at Nestlé noticeably damaged the brand when a consumer post about the destruction caused by palm oil forestation was answered by a belligerent company representative, ultimately transporting the resulting discussion to mainstream media [24].

In order to provide insights to practitioners looking to use social media to benefit their brands, scholars have focused on users by trying to identify the most influential target group [55] or explain their relation to the social media [56]. Others have addressed the challenges of social media marketing such as aggressive advertisement and over-commercialization, lack of e-commerce abilities, invasion of user privacy and transparency control [8], [57]. An inappropriate approach to these challenges could lead to fan loss and exposing the company to the risk of destroying its own credibility. In turn, a response to the negative feedback which complies with the collaboration demands of the social media platform users might convert the critics into active and enthusiastic supporters of the brand [55]. As an example, General Motors responded to the criticism with the invitation of these people to visit the factory where addressed problems were explained together with the undertaken solutions resulting in a positive outcome.

Apart from the challenges, many opportunities have also been recognized. According to Richter et al. [40], social media platforms offer the potential for (1) advertising - by facilitating viral marketing, (2) product development - by involving consumers in the design process and (3) market intelligence - by observing and analyzing the UGC. In addition, engagement on social media platforms could raise the public awareness about the company, enable community involvement and provide support for gathering experience for the future steps [8]. Finally, as Javitch [31] advises, free social media marketing is a good alternative to the costly traditional marketing campaigns and getting involved on SNs also means protecting the business name.

Although many social media marketing channels have already been created, how these channels are being used, what their potentials are and how consumers interact, remains largely unknown. A structured, academic analysis is still outstanding and has yet to be addressed from different perspectives [40].

3. Brand Presence on Social Media

Based on the analysis of the existing examples of brand presence on social media, we propose the following classification of the modes of interaction between the companies and the consumers:
• **Brand pages** represent profile pages created by the companies, offering the opportunity for active involvement, both, on the side of the brand owner, as well as the customer, by engaging in a direct communication.

• **Applications** provide the possibility for an additional approach towards advertising, by building upon the social concepts of the underlying platform, in turn offering benefits and convenience to the customers.

• **Social plugins** allow simple integration of social media concepts into the existing online and mobile platforms, providing possibility for targeted marketing to the companies, and direct insights into their friends’ opinions to the consumers.

• **Public content** differs from the previous concepts in a way that it does not represent a specific tool or a platform. In this paper, the term will be used to refer to the totality of publicly available content, shared by the users on their profile pages.

In the continuation, each of these concepts will be explained in details focusing on the format of the available data and the provided level of details.

### 3.1. Brand Pages

Brand page is a term used to describe profiles created on social media platforms by brand owners [48]. Brand pages represent a natural technological platform for marketing, providing access to a large number of users, grouped in communities and based on a structured set of social relationships among admirers of a brand, i.e. a brand community [49]. Depending on the communication policy set by the company, consumers might have the possibility to share content such as status updates, photos, videos, etc., or they might be able to comment on the content created by the company. In addition, interactions in a form of “liking” and “sharing” are usually supported on these platforms.

Figure 1 illustrates an example of a Facebook brand page showing the content shared by the users, i.e. “page fans”.


Brand pages are a valuable source of UGC which might reveal topics that attract the attention of the consumers, their intentions for participation and emotions shared over certain product or brand [50]. As such they provide possibility for collection of: (1) quantitative data i.e. number of likes, comments, and shares; and (2) qualitative data, i.e. product/brand related opinion contained in the status updates and/or comments. In addition, brand pages usually provide summarized demographics information over the page fans, such as age and gender distribution, language and location, etc. Although this information can’t be related to a particular product, it can be valuable for the brand owners to gain understanding over the overall demographics of their target audience, as well as the possible changes caused by specific marketing campaigns.

UGC created on brand pages introduces challenges for the text mining applications. Since these platforms do not pose any limitations into the form and topics reflected in the status updates, extracting the information regarding the targeted product might be a challenge due to the informal language used on social media platforms [43].

### 3.2. Applications

The concept of social application is supported so far only by Facebook. Integrating an application within the Facebook platform provides the opportunity for accessing the core Facebook components while maintaining the same social experience into the intended service. Today there are many applications on Facebook, with the social games leading on the leader boards [3].
The popularity of a Facebook application measured by the number of users has attracted the attention of companies and service providers who have started utilizing different approaches in order to gain knowledge or increase the size of their community. The possible gain from this approach depends solely on company’s creativity to exploit the benefits of the common Facebook interaction possibilities, such as liking, sharing, commenting, etc. This in turn could result into the well structured UGC that could be directly linked to the targeted product. Both, quantitative (likes, comments, shares) and qualitative data (product related comments) could be obtained from this source of information.

Additional information available through Facebook applications is demographics information of the users. While brand pages, public posts and social plugins provide very limited information into the users demographics on individual level (e.g. Facebook allows insight only info the gender information for non-public profiles), applications usually ask users to grant them consent to access the private data, such as age, language, number of friends, etc. This might bring additional insights to the companies in terms of understanding the appeal of specific products to particular demographics group.

Recent trend in Facebook applications goes into direction of creating platforms for online shopping on Facebook, i.e. F-commerce [18]. An example of such application is illustrated on Figure 2.

Fig. 2 F-commerce application as an online shopping platform on Facebook.

F-commerce is still in its infancy, although early attempts can be traced back to 2007, when Facebook tried the Project Beacon which collected e-commerce activities on third party sites
and announced a user’s purchase actions on his or her friends’ news feed. This attempt was soon declared as a failure due to the mass disapproval from the users related to the privacy issues [39]. Yet, recent studies showed that 48% of the millennials (aged 20-33) would like to see their favorite stores providing them with the possibility to buy directly on Facebook [20].

3.3. Social Plugins

According to the definition provided by Facebook, “social plugins let you see what your friends have liked, commented on or shared on sites across the web.” [22] The idea behind the social plugins is to provide the possibility for simple integration of the existing online platforms with the social media.

It all started in 2008, when the Facebook Connect was launched allowing users to sign in to external websites using their Facebook accounts. This approach soon became very successful, reaching 100 million users on Web and Mobile sites in just one year [42]. The potential of social plugins was soon recognized by other social media sites, resulting in the announcement of similar plugins, e.g. the “Tweet” button in August, 2010 [46]. Today, most of the biggest social media platform support some form of social media plugins which allow expressing opinion, commenting, sharing, recommending, etc., as illustrated on Figure 3.

Fig. 3 An example of usage of the “sharing” social plugin.

The potential of social plugins for the marketers has already been recognized as an opportunity to increase the engagement level and thus deliver specific marketing messages to the optimum target market [1]. Social plugins provide the possibility to collect (1) quantitative data i.e. number of likes, shares and recommendations and fans, and (2) qualitative data, i.e. comments written by the users which reveal their opinions. It should be noted that when it comes to the number of fans, some platforms require ‘liking’ or joining the page in order to provide the users with the possibility to post a comment which might also be a negative one. Therefore, companies should be careful when interpreting the number of fans and should take in consideration the specific requirements of the underlying platform.

Although social plugins can be used to automatically generate fans, i.e. to add members to the brand community (for example by pressing the Facebook “Like” button), the real value for extracting knowledge originates from the core concept behind the plugins, i.e. integration should be performed on a level of profile pages representing real-world things, such as products. This would provide the possibility to clearly distinguish between opinions over individual products, without the need to perform complex text mining operations.
3.4. Public Posts

Public posts are the most general form of UGC not related to any specific mode of social media platform utilization. Instead public posts correspond to the status updates, uploaded videos, photos, locations, or other types of content shared by those social media users who have set their privacy policy to “public”. In other words, this is the content generated by the users who have made their actions visible to all the other user of the underlying platform.

Public posts represent unstructured content and can be shared in various formats. In order to extract product related information, a company must search for brand/product mentions to extract the relevant subset. Most of the social media platforms, e.g. Facebook, Twitter, etc., provide interfaces and APIs to support this mode of utilization [51], [52], [53]. As a result, in the domain of social media marketing, there are numerous commercial tools for monitoring and tracking of social media brand presence over multiple brand channels [34]. This process is commonly known as social media monitoring.

In addition, public posts represent valuable source that provides insights into the topics that attract large fraction of users [54]. Automatic extraction of such topics is known as trend detection and monitoring and for the companies it provides the possibility for benchmarking against competitors.

Regardless of the form of social media brand presence or the source of products and brand related content, tools that provide the possibility for automatic information retrieval are needed in order to enable efficient knowledge extraction over the vast amount of UGC on social media platforms.

4. Extracting Knowledge from the User Generated Content

The value of the UGC as a source of information was already recognized, resulting in individuals turning to social media platforms as sources of real-time news and opinions [41], [33]. Public opinions shared on social media platforms are interesting not only for individuals, but also for (1) news reporters [17], pointing to the fast-evolving news stories, (2) sociologists, revealing the ‘spirit of the times’ [37], (3) opinion tracking companies, e.g. for prediction of elections outcome [45], and (4) marketing professionals, for brand image monitoring and benchmarking [30].

This form of usage of social media platforms has further been supported by the platform providers, by offering the possibility for searching through the public status updates to monitor content or find temporally relevant information [44]. In addition, they have offered the possibility to access the public status updates through their APIs [52], [53], resulting in a
burst of commercial and research efforts to gather knowledge through analysis of the shared content.

As the number of available sources and the amount of online information increase, individuals and companies interested in gathering knowledge through monitoring of the conversation on social media platforms need to rely on the tools capable of automatic document analysis and classification. This is generally known as information retrieval.

4.1. Information Retrieval

Information retrieval is a research field related to the problem of finding information or content from a corpus of documents which corresponds to some more-or-less well-specified query. The term was first introduced in 1950 by Mooers, as “the problem of directing a user to stored information, some of which may be unknown to him.” [36] Since then, the field of information retrieval has grown beyond indexing and searching. Today, it involves the problems of modelling, document classification and categorization, system architecture, user interfaces, data visualisation, etc.

The major tasks to be accomplished when utilizing the information retrieval techniques are:

- representation of documents,
- representation of queries, and
- techniques for comparison of documents and queries [6].

One common approach towards overcoming these challenges is transformation of the documents into a set of index terms as represented on Figure 4.

![Fig. 4 Document transformation: from full text to a set of index terms [5].](image)

This approach is applicable to the content shared on social media. Each status update could be stored in a database and further transformed to a set of index terms. Extraction of product related data could then be reduced to keyword based search over index terms, where keywords would be brand and/or product name.
It should be taken in consideration that the change brought about by social media towards short commentary, as introduced by SNs such as Twitter and Facebook, resulted in a significant difference in the comment structure and language, imposing challenges to the existing text mining techniques [43]. Further research and improvements of existing methods would benefit towards the possibility to fully harvest the value of the UGC on social media platforms.

### 4.2. Opinion Mining and Sentiment Analysis

Related to the understanding of the conversation on social media platforms are also research fields of opinion mining and sentiment analysis. Insights obtained through opinion mining and sentiment analysis allows companies to determine how the population perceives a given brand, product or feature, i.e. for market analysis and rumour detection.

The term opinion mining was coined by Dave et al. to describe a tool that would “process a set of search results for a given item, generating a list of product attributes (quality, features, etc.) and aggregating opinions about each of them (poor, mixed, good).” [16] Approximately at the same time the term sentiment was used in reference to the automatic analysis of evaluative text and the tracking of predictive judgments [15]; thus the two disciplines are interconnected.

In the most general sense, sentiment analysis refers to the process of polarity classification into positive, negative or neutral. Today, there are existing tools, such as LingPipe [35], based on machine learning techniques to automate sentiment analysis over large collections of documents. The underlying concept of the machine learning techniques is training the classification algorithm based on manually classified dataset, as illustrated on Figure 5.

![Fig. 5 Sentiment classification based on machine learning.](image)
An additional approach is sentiment classification based on semantic lexicons, such as SentiWordNet [21], in which each term is already assigned with a sentiment score previously derived from the semi-supervised classification.

It should be noted that despite the maturity of the field, sentiment extraction is not a straightforward process. While computers can easily detect existence of a word representing sentiment within a text, interpretation of slight variations in the tone, such as sarcasm or slang, which can dramatically affect meaning, is not mastered yet. Since this task is difficult even for humans, sentiment analysis will probably remain an imperfect discipline.

5. Integration of Social Media into the GS1 Framework

In parallel to the growth of social media, with over four billion [12] users in the world, smartphones are becoming an additional component in the online communication between the companies and consumers. Smartphones are becoming sensing devices which enable context recognition based on the data obtained from the built-in sensors. Among the possibilities offered by their current technological features is product recognition based on barcode or RFID identification which in turn enables closer interaction between the consumers and their favourite products and brands [28].

With the continuous growth of the number of smartphone users, brand owners are being faced with new challenge: How to utilise smartphone technologies to effectively connect and engage with consumers? As a result, many retail brands have focused on integration with social media platforms such as Facebook, Twitter, and Foursquare, not only to create better shopping experiences for their customers, but also to take advantage of the fact that by
social media integration they might increase the overall brand awareness. This assumption is based on the fact that the likelihood of social interactions, such as “sharing” and “liking”, might also increase by providing appropriate service, which in turn would increase the brand visibility in the social space. This would lead towards achievement of the companies’ goal across multiple social media platforms.

The opportunities offered by the utilization of smartphones for providing on-site product information to the consumers were already recognized by the GS1 organization [26]. The proposed integration of the smartphones with the existing GS1 framework components is known as MobileCom/B2C solution [27].

MobileCom/B2C is based on the reasoning that consumers are becoming more demanding in regard to the information provided to them, in particular in the level of details of provided information and immediate and on-site accessibility. With this in mind, MobileCom/B2C architecture proposes the following interaction steps between the GS1 framework and the consumers [2]:

- Steps 1 and 2: The product identifier (GTIN) is decoded from the barcode and passed to the Object Naming Service (ONS).
- Step 3: The ONS responds by providing the location of the information in the relevant data aggregator.
- Step 4 and 5: A request is passed to the aggregator and the information for the requested product is returned.
- Step 6: This information is rendered by the application for the consumer.

Starting from the existing MobileCom/B2C architecture [2], integration with social media would be as simple as adding an additional content provider block, i.e. the Social Media Content Provider (SMCP). Since the content generated on social media should be presented to the user at the same time and through the same interface as the product information, this means that the aggregator should be capable of communicating with the SMCP in order to obtain the “social content” for the requested product.

As a first step, an identifier should be sent to the SMCP to describe the required information. Since SMCP is not familiar with the format of the GTIN, the content retrieval from SMCP should be based on the product information retrieved by the aggregator, i.e. brand/company name and/or product name. Thus an additional step would be added in the current process:

- Step 4.1: Aggregator accesses the SMCP based on the retrieved product name and obtains the product related UGC.

The resulting framework architecture is illustrated on Figure 7, showing the proposed integration of the social media into the existing MobileCom/B2C GS1 architecture as well as the new level of details provided to the consumer.
Fig. 7 Integration of Social Media into the GS1 MobileCom/B2C architecture.

It should be taken in consideration that, as already discussed, social media provides large volume of unstructured data. For that reason, social media content provider should be capable of satisfying the previously described requirements:

- Data collection from the relevant source, i.e. social media platforms,
- Information retrieval, i.e. extraction of product related content based on product/brand name query,
- Sentiment analysis over the filtered dataset, and
- Unique data representation for different formats of content shared on social media platforms.

As a result, apart from obtaining information such as nutritive values and price, consumers will be able to access their friend’s opinions, as well as the overall opinion of the online communities from different social media platforms by looking at the:

- Quantitative information:
  - Product/brand related status posts, and
  - Product/brand related comments;
but also at the summarized information in form of:

- Quantitative data, such as:
  - Number of posts/mentions,
  - Number of comments,
  - Number of positive vs. negative posts/comments,
  - Number of likes,
  - Number of shares, and
  - Number of recommendations.

To summarize this discussion, we propose the architecture of a framework for extraction of product related information from the UGC from different social media platforms. The graphical representation of this framework is illustrated on Figure 8.

We build our proposal based on the discussion presented in previous sections. First, we showed that social media platforms provide four different possibilities for creation of brand presence. We than explained the differences in the level of details of the UGC for each of these platforms and challenges that originate from these differences. Finally, we provided an overview of the existing text mining techniques that provide the possibility to extract and quantify the product related information.

The proposed framework architecture should be used as a starting point for design and implementation of a social media content provider.
6. Conclusions and Future Steps

In this paper we provided a solution for integration of social media into the GS1 architecture. We first explained the growing importance of the social media with a focus on social media marketing. We then presented different possibilities for creation of social media brand presence, and the tools that could be employed for information extraction from the UGC on these platforms.

We believe that the opportunities that would emerge from the proposed integration could be taken as a clear indication of the validity of the presented solution. These would go in two directions, providing benefits to both, brand owners and the consumers.

As discussed in section 2, opportunities for social media integration for the brand owners have already been recognized. These include:

- Conducting brand monitoring,
- Gathering direct insight into the consumers’ opinions, by
- Avoiding high expenses related to traditional market research,
- Generation of ideas for new products,
- Benchmarking to the competitors, etc.

At the same time, consumers would be able to gather insights into the opinions of their closest friends, but also of the online community in total. Various data representations and additional sources of information could support the users into achieving different goals, e.g. finding out what their friends think about the chocolate cake in Starbucks, but also what is the most popular Starbucks location in the city.

To further contribute in this direction and explore additional possibilities, such as for example the integration of the location as illustrated above, we plan to continue our work by providing practical implementation of the proposed concept and evaluation of different possibilities for information retrieval and presentation to the users. Finally, we would like to explore the possibilities of expanding the number of information sources beyond the social networks by utilizing new concepts aiming at accessing and collecting the data from the web in general, such as OpenID, Friend-of-a-Friend (FOAF) ontology, etc.

References


