Corporate Communications: An International Journal

CSR communication strategies for organizational legitimacy in social media
Elanor Colleoni,

Article information:

Users who downloaded this article also downloaded:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

CSR communication strategies for organizational legitimacy in social media

Elanor Colleoni

Department of Intercultural Communication and Management,
Copenhagen Business School, Copenhagen, Denmark

Abstract

Purpose – Organization legitimacy is a general reflection of the relationship between an organization and its environment. By adopting an institutional approach and defining moral legitimacy as "a positive normative evaluation of the organization and its activities", the goal of this paper is to investigate which corporate communication strategy adopted in online social media is more effective to create convergence between corporations' corporate social responsibility (CSR) agenda and stakeholders' social expectations, and thereby, to increase corporate legitimacy.

Design/methodology/approach – Using the entire Twitter social graph, a network analysis was carried out to study the structural properties of the CSR community, such as the level of reciprocity, and advanced data mining techniques, i.e. topic and sentiment analysis, were carried out to investigate the communication dynamics.

Findings – Evidence was found that neither the engaging nor the information strategies lead to alignment. The assumption of the more the dialog, the more the communality seems to fail to portray the complexity of the communicational dynamics, such as the persistence of different, or simply a dialog without alignment. Empirical findings show that, even when engaging in a dialogue, communication in social media is still conceived as a marketing practice to convey messages about companies.

Originality/value – This paper originally investigates organizational legitimacy in the context of social media by applying advanced data-mining techniques that allow the analysis of large amounts of information available online.

Keywords Organizational legitimacy, Communication strategies, Corporate social responsibility, Social media, Topic analysis, Sentiment analysis, Social responsibility, Communications

Paper type Research paper

Introduction

Corporate legitimacy has become a pressing issue inasmuch as the stakeholders' perception of the role of business in society has been significantly redefined. New ethical expectations have risen along with a set of contingent social responsibilities that corporations are now asked to fulfill by the various groups of stakeholders in society (Moreno and Capriotti, 2009).

As pointed out by Scherer and Palazzo (2007), these growing expectations are explained in part by the process of globalization, which is eroding established national institutions and procedures of governance. This particularly applies to corporations operating at a global level, where international standards of production, as well as labor regulations, often fail to constrain their behaviors and leave them the responsibility to self-regulate their production processes. Furthermore, during the last

few years several companies have been involved in social and environmental disasters and, as a result their legitimacy has been challenged (Palazzo and Scherer, 2006). As a consequence, citizens are increasingly demanding that the corporations justify and legitimate not only their economic actions, but also their social and environmental actions in the general public sphere. Redefining the relationship between business and society brings about the creation of a new form of corporate legitimacy and implications for businesses to open a dialogue in society. Corporate legitimacy has been defined as "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions" (Suchman, 1995, p. 574). The creation of a "congruence between the social values associated with or implied by [organizational] activities and the norms of acceptable behavior in the larger social system" (Dowling and Pfeffer, 1975, p. 122) lies at the core of the legitimacy of business in society. Eventually, obtaining legitimacy by aligning corporate behavior with stakeholder expectations is necessary to guarantee the corporation's continued existence (Dawkins, 2004).

Following this argument, corporate social responsibility activities encompass all of the diverse corporate social practices implemented in order to increase the congruence between corporate behavior and the social expectations of stakeholders and, as a consequence, to increase their legitimacy within society. Several researchers have shown the relationship between a firm's engagement with CSR and its economic performance, the well-known "doing well by doing good" argument (Bhattacharya and Sen, 2004; Wood, 1991).

However, in order to be perceived as socially responsible organizations, companies must match stakeholders' social expectations, which may not always be an easy task. Indeed, stakeholders' evaluations of CSR activities are not always positive; for instance, if those activities are perceived as having been achieved at the expense of product quality or if it is perceived that CSR efforts do not enhance corporate abilities (Morsing *et al.*, 2008; Bhattacharya and Sen, 2004). As pointed out by Dawkins (2004), the ability to establish congruence between social expectations and CSR agenda depends on the capacity to communicate with different stakeholders and on the support obtained from audiences; moreover, the difficulty is "how to make it known [...] and how the company deliberately should communicate it" (Morsing *et al.*, 2008, p. 98).

Stakeholder awareness has been dramatically enhanced by the advent of the internet, which enables people to share almost anywhere with almost anyone "connected" on a scale that has not been seen in the past (Colleoni *et al.*, 2011). Furthermore, with the diffusion of social media, such as Twitter and Facebook, stakeholders are no longer passive receivers of the communication, but increasingly engage in the creation and evaluation of content (Dellarocas, 2003). These features represent both great potential and a challenge for corporations; a strategic plan on what to present and how to present their CSR commitment to stakeholders is becoming a pressing issue (Dawkins, 2004).

Due to corporate fear of potential stakeholders' skepticism, online communication channels, regardless of their specific features, are often used as non-interactive traditional mass media for corporate self-presentation (Insch, 2008). On the other hand, the evolution of Web 2.0 allows corporations to easily and directly engage with stakeholders in a dialogue, which can potentially bring co-creation of meaning: It also allows the emergence of a joint setting of the corporate CSR agenda and shared social values (Kent and Taylor, 1998).

Although these communicational strategies have received substantial attention (Capriotti and Moreno, 2007; Morsing and Schultz, 2006; Dawkins, 2004; Ryan, 2003; Schultz, 2000; Kent and Taylor, 1998; Botan, 1997; Rafaeli and Sudweeks, 1997), there is still a lack of systematic knowledge regarding whether and in which measure these different strategies are able to create a bridge between the stakeholders' social perceptions of corporate responsibilities and the corporate CSR agenda, and consequently, to increase corporate legitimacy. This study investigates the CSR network community on Twitter, an interactive micro-blogging platform, and its actors, i.e. corporations, audiences and experts. It analyzes empirically which are the communication strategies adopted by the corporations and whether these strategies lead to the matching of corporate CSR agenda with stakeholders' social expectations.

The key contribution of this paper lies in the empirical evaluation of which CSR communication practice is more effective in creating congruence between corporate and stakeholders' social values.

This paper is organized as follows. In the first section, the relevance and peculiarity of new social media sites as communication channels for reaching stakeholders is introduced and the different communication strategies that can be adopted are discussed in detail. A theoretical framework that summarizes the different strategies and their implications is also presented. The second section is devoted to operationalizing the framework empirically in its various components. Firstly, in the third section, the resulting communication strategies empirically identified in the network are described along with the topics of discussion and their affective orientation. Secondly, the matching output of the different communication strategies, in terms of opinion similarity and dissimilarity between stakeholders and corporations are presented and theoretically interpreted as an indicator of congruence between social expectations and CSR corporate agenda. Finally, the theoretical framework is discussed in light of the results along with suggestions for further research in the field.

Theoretical background

The concept of organizational legitimacy stresses the fact that "many dynamics in the organizational environment stem not from technological or material imperatives, but rather from cultural norms, symbols, beliefs, and rituals" (Suchman, 1995, p. 571). Organizational legitimacy represents the conformation to these values, norms and expectations that are socially constructed within society (Berger and Luckman, 1966). This conformation is built by recognizing the necessity to interact with all of the groups, inside and outside the corporate environment, that have interests in the activities of the organization (Freeman, 1984). Legitimacy is vital for organization survival (D'Aunno and Zuckerman, 1987). Cultural congruence empowers organizations primarily "by making them seem natural and meaningful and also more trustworthy" (Suchman, 1995, p. 576). Therefore, companies perceived as desirable, proper or appropriate are more likely to be institutionalized and, therefore, to be persistent (Parsons, 1960). Furthermore, organizational legitimacy can be a goodwill buffer during adverse times (Suchman, 1995). By contrast, organizations with frail legitimacy are more subject to the risk of being perceived as unnecessary or irrational (Dowling and Pfeffer, 1975). Legitimacy is managed successfully when "organization actions are perceived as being consistent with stakeholders' expectations" (Massey, 2001, p. 156). Suchman (1995) has identified three types of legitimacy:

(1) pragmatic;
(2) cognitive; and
(3) moral.

Pragmatic legitimacy will be conferred by the audiences through direct exchanges and based on their self-interests and determination of practical consequences (Suchman, 1995). Cognitive legitimacy is reached when an entity becomes embedded in taken-for-granted assumptions (Zucker, 1987). Moral legitimacy reflects a positive normative evaluation of the organization and its activities (Aldrich and Fiol, 1994). In this latter case, the audience evaluates organizations not based on their immediate benefits, but on whether it is "the right thing to do" (Suchman, 1995, p. 579). The globalization process, along with the loss of national power, the emergence of the risk society and general economic social and environmental uncertainty (Beck, 2000; Castells, 1997) has increased the importance of reaching moral legitimacy (Herrmann, 2004). Corporations are seen as actors that must act for the general wealth of society and are increasingly treated as personified autonomous subjects. Although corporations are progressively translating their corporate values into tangibles corporate social responsibility practices, a correspondence of values between stakeholders and corporate is not guaranteed; a mechanism of communication and feedback must be established with the audience and society as a whole. This implies that an organization must develop a sense of organizational values to be communicated and thereafter be willing to revise its conception based on the influence others (Morsing and Schultz, 2006). Meyer and Rowan (1977) first called attention "to the ways in which organizations seek legitimacy and support by incorporating structures and procedures that match widely accepted cultural models embodying common beliefs and knowledge systems" (Meyer and Rowan, 1977, p. 878). The definition of these accepted cultural models has been traditionally conceived as the result of an internal process of understanding and shaping of the corporate environment. In contrast, the evolution of the stakeholder theory has increasingly empathized how the link with the external environment must be conceived, not only as a flow of information, but as a communication channel through which an organization can co-create meanings with its stakeholders (Andriof and Waddock, 2002). These two different organizational communication approaches have been extensively investigated by corporate communication. Several authors have found evidence of these two antithetical communicational approaches. They have been named in different ways: the transmission and the ritual view (Carey, 1989); and the monologic and the dialogic view (Ingenhoff and Koelling, 2009; Ryan, 2003; Kent and Taylor, 1998; Botan, 1997). Based on Grunig and Hunt's public information model (1984), Morsing and Schultz (2006) have developed a model based on two dimensions, the interpretation of corporate values and the integration of stakeholder feedback, and they also identified three main communication strategies: information, response and involvement. It is the goal of the information strategy to disseminate CSR corporate information. Here, the interpretation of corporate values and the resulting CSR agenda are mainly defined internally, while the external feedback is not integrated into the corporate vision. In the response strategy, the definition of corporate values is still an internal practice, but is addressed by monitoring stakeholder reactions to CSR activities. Based on an empirical analysis of Danish companies, Morsing *et al.* (2008) have redefined and divided this

strategy into two different strategies: the expert and the stakeholder endorsement strategies, which are used to diffuse content and get feedback while avoiding potential stakeholder criticism. In this case, "companies target [. . .] an exclusive group of experts [. . .] which the companies themselves regard as 'elite readers' of corporate messages" (Morsing *et al.*, 2008, p. 105), or they target a specific group of stakeholders and rely on them to convey the corporate CSR messages. The involvement strategy implies no boundaries for the organization that co-creates meaning, sense and vision together with its stakeholders. This dialogical relationship affects both the definition of corporate values and its CSR practices. In Figure 1, I present the theoretical framework, accounting for the relationship between corporate communication strategies and legitimacy where the congruence level (i.e. the match) is unknown.

In this framework, legitimacy is theoretically defined as the congruence between stakeholders' social expectations and corporate CSR agenda and empirically described as the quality of the match in terms of similarity of attention and affective orientation towards CSR related themes. Eventually, the quality of the match is the assessment of the congruence of values, which begins and ends with the two main actors and their ethical choices (i.e. CSR agenda and social expectations). Indeed, since I do not assume different a priori convergence between the communicational strategies, the degree of the match and the level of legitimacy associated must be verified after the communication process. The framework outlines three main communication strategies based in part on Morsing and Schultz's (2006) model:

(1) self-centered;

(2) mediated; and

(3) dialogical.

The self-centered communication strategy implies that companies define their CSR agenda internally. Unlike Morsing and Schultz (2006), who outline two distinct strategies based on internal CSR agenda definition – one that benefits from external feedback (i.e. response strategy) and one that does not (i.e. information strategy) – I assume that companies always have the opportunity to benefit from feedback and to converge stakeholders' expectations. In this case, the congruence will be the result of an internal process based on feedback about corporate legitimacy. In the mediated
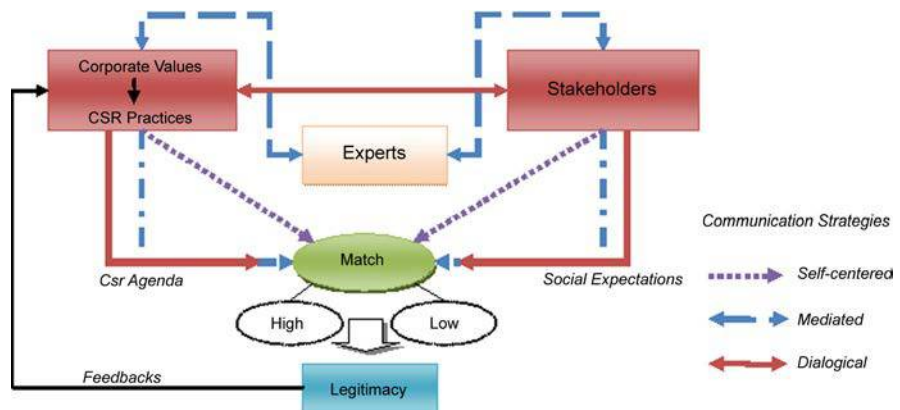

Figure 1.
Communication strategies for corporate legitimacy

communication strategy, the central actors are the experts who bridge corporations and stakeholders by endorsing companies and establishing a dialogue with stakeholders in order to convey third-part messages. In the dialogical strategy, the central actors are the stakeholders who actively interact with the corporations, and the alignment results from direct feedback (Morsing and Schultz, 2006).

The goal of this paper is to investigate whether these different communication strategies lead to congruence between stakeholders' social expectations and CSR agenda.

Online communication in social media

In order to investigate the corporate communication practices and their impact in terms of legitimacy, an analysis of online corporate communication in social media has been carried out. Indeed, online communication is becoming one of the most important information channels for corporations aiming to increase stakeholders' awareness about their CSR efforts. The internet has reshaped communication at different levels. Indeed, traditional broadcast media are based on a hierarchical one-to-many communication, with a clear distinction between producer and consumer of information and an audience that does not participate in the creation and selection of content. In contrast, The internet has enabled the emergence of a new participatory public sphere based on a many-to-many communication where everybody can dialogically and publicly interact and collaborate in the creation of content and the definition of the agenda (Jenkins, 2006). For instance, www.amazon.com, the world's largest online retailer, has based its success on the trust generated by its customer product reviews. With the diffusion of social media such as Twitter and Facebook, attracting people and creating an active community is no longer a difficult and expensive task for companies. However, since the relationships are organized around a network of peers with no centralized entities of control in social media, corporations are thrown into a global stream of communication as one of the stakeholders. Within these media communication is referred as "viral" because ideas and opinions spread like epidemic diseases through the network via word-of-mouth and are perceived as highly trusted by the users because they are based on group similarities (Colleoni *et al.*, 2011). In the analysis I investigated social media sites, particularly Twitter, where communication is horizontal (i.e. democratic) and corporations can only share their own agenda, and thus the emergence and the similarity of topics can be regarded as the result of network communication dynamics.

Research questions

I outlined the following research question based on the theoretical framework:

RQ1. What are the characteristics of CSR communication in social media?

Particularly, I am interested in investigating the communication strategies adopted by the corporations in an open environment. In which type of environment do they operate? Do they tend to share the public with other companies or do they have specific audiences? Have corporations have adopted an interactive approach with their stakeholders, or only with some of them?

RQ2. Is there an alignment between companies CSR agenda and stakeholders' social expectations in social media? If there is alignment, do the different communication strategies lead to different alignment outcomes?

With the introduction of new media, and particularly interactive social media, a dialogical conception of the relationship between an organization and its stakeholders has gained increased attention (Andriof and Waddock, 2002; Palazzo and Scherer, 2006). The attractiveness of the relational approach is closely related to the emergence of a new informational society and the stakeholders' empowerment that this has entailed (Castells, 1997). Stakeholders' perceptions of corporate behavior are built more independently from a companies' information and traditional mass media and they increasingly rely on direct communication within networks of peers. This transformation has posed the condition for the emergence of a new public sphere (Latour and Weibel, 2005). Following Habermas's (1989) theory of public sphere, the general assumption is that if the corporation directly engages with stakeholders in the definition of the CSR agenda, the communication process will automatically generate the emergence of a common interpretation of reality (Smith, 2003; Morsing and Schultz, 2006; Palazzo and Scherer, 2006; Basu and Palazzo, 2008). However, this assumption has never been systematically investigated. It can be argued that establishing a communication does not guarantee the creation of a dialogue (Lyotard, 1984). For instance, the interactive communication is often more devoted to persuading others rather than building a common view (Coupland, 2005); moreover, corporations might operate a selection in the topics of discussion and avoid debating controversial issues (Schultz et al., 2011).

Methodology
Identifying Twitter CSR community and its actors
In order to investigate companies' strategies, I have investigated communication practices on Twitter. Twitter is a micro-blogging service launched in 2006 that allows users to describe their current status via short posts (i.e. "tweets"). Using the entire Twitter social graph from June to December 2009, I selected the companies from the 100 Best Corporate Citizens 2009 list redacted by *Corporate Responsibility Magazine* (2009) who have a Twitter account dedicated to corporate social responsibility. I found seven CSR corporate accounts: Starbucks, Ford, Microsoft, Campbell, Smarter Planet, Xcel Energy and Green Mountain Coffee Roasters. I detected the social graphs of each CSR account (i.e. the list of followers and following) and I identified a first CSR community consisting of 16,000 users. However, I filtered this initial network in order to extract only CSR-related content from the flow of communication and those users who have shared CSR at least once content within the network. In order to do so, I first extracted a sample of the conversations and manually coded 3,000 tweets according to the distinction CSR/non-CSR content. I then used this sample as a training database to create an automated classifier able to detect CSR content in the whole network. Using the training database and a learning algorithm, the classification algorithm learnt how to classify CSR tweets, reaching an accuracy of 89.33 percent of correctly classified tweets (see Appendix 1). Therefore, I was confident to apply the classifier to the entire set of tweets. Once all the conversations were classified by CSR content, I could identify and select those users who had shared at least one CSR-related tweet within the community. I ended up with a network of 9,589 users and 326,000 CSR-related tweets, consisting of two groups, i.e. the corporations and their followers. While corporations are represented by CSR accounts, the followers are divided in two groups, i.e. audience and experts. I label the audience as people who follow one corporate account and I

assume that they have a specific interest in the company they are following. In contrast, I call the CSR experts users who follow more than one company, showing a general interest in CSR.

Corporate communication strategies
In order to identify the corporate communication strategies empirically, I had to translate the conceptual artifacts into measurable Twitter communicational practices. In the following section, the three general communication strategies (i.e. self-centered, mediated and dialogical strategies) are categorized according to two dimensions:

(1) the network structure, i.e. how corporations create their relational social ties; and

(2) the communication flow, i.e. how corporations produce and share content within their own network.

The first dimension accounts for the different configurations of the relational ties in the CSR community, i.e. the following/followers relationships. Indeed, in contrast to Facebook, which is based on a two-way friendship and connects people already known in real life (i.e. social graph) Twitter is in part also an interest graph, i.e. a direct network based on both one-way following and two-way friendships (Ravikant and Rifkin, 2010). This peculiarity allows for "micro-celebrity" and strategic personal branding practices aimed at reaching prestige within the network (Marwick and Boyd, 2011). For instance, Microsoft founder Bill Gates follows only 71 people while being followed by more than three million; in contrast, US President Barack Obama follows almost one million people and is followed by ten million. In the analysis, I considered the mutual recognition of a relationship and the creation of closer ties in the network by calculating the degree of reciprocity. Reciprocity indicates the degree of cohesion, trust, and social capital that is present in a network and is measured as the tendency of node pairs to form mutual connections between each other (Borgatti, 2005). According to the framework, different communication strategies are associated with different levels of reciprocity: in the case of the self-centered strategy, the communication channel is used as a one-directional tool; I therefore expect a low level of reciprocity. In the mediated strategy, the experts are trusted by the company as endorsers of their communication; I therefore expect the company to establish a higher degree of reciprocity with them rather than with its stakeholders. In the dialogical strategy, using public dialog, the company should act as a facilitator for the creation of a connected community; I therefore expect a high degree of reciprocity.

The second dimension accounts for the different communicational practices among the actors. Besides the personal broadcasting function, Twitter allows users to interact dialogically and share content. Direct messages represent the main mechanism of interaction in Twitter and are used to both address a message to a specific user, or reply to someone's tweet always in a public context, so that "when conversations emerge, they are often experienced by broader audiences than just the interlocutors" (Boyd *et al.*, 2010:p.1). Retweeting is the basic mechanism through which users share content and it can be seen as one of the important means of opinion formation. This practice "contributes to a conversational ecology in which conversations are composed of a public interplay of voices that give rise to an emotional sense of shared conversational context" (Boyd *et al.*, 2010, p. 1). According to the theoretical model,

companies adopting the self-centered strategy will make low or no use of direct messages and retweet. The mediated strategy describes a situation where companies address directly and retweets experts' messages. The dialogical communication strategy accounts for those companies who interact profusely with their audience via direct/reply messages and who forward their audiences' content to their own public, facilitating dialogue among stakeholders. To summarize, in the self-centered strategy, the company does not reciprocate the social connection and does not retweet or directly refer to its followers' content. In the mediated strategy, the creation of social ties and CSR communication between the company and the audience is mediated by the CSR experts. In the dialogic strategy, the company creates a mutual connection with its audience and establishes a direct dialog.

Attention and affective orientation towards topics

CSR discourse can be seen as a set of themes and their normative goals. However, the idea of what is good and what is important may differ among actors. I considered the relative attention and the affective orientation towards topics as a proxy of what is considered "important and good". Reaching legitimacy requires the overlap of these two elements among stakeholders and companies. In order to investigate CSR discourse among groups empirically, I first extracted the different themes addressed by the actors in the network. In order to identify topics of discussion, I applied automatic topic detection using latent semantic analysis (LSA; see Appendix 2). It is important to stress that LSA does not provide an optimal number of topics. This choice must be made by the researcher according to the theoretical framework. After having identified the topics, I defined the attention and the affective orientation for each group (i.e. corporate, audience and experts) by topic. I measured the relative attention around a topic as the rate of tweets referring to a particular topic divided by the total number of tweets.

I used sentiment analysis to measure the affective orientation towards a topic. Sentiment analysis refers to the process of categorization of unstructured human-authored documents "based on their affective orientation" (Colleoni *et al.*, 2011). Indeed, with the availability of massive data produced in social media sites and a set of new algorithms having been developed, it has become possible to deploy data mining techniques such as opinion mining and sentiment analysis in order to monitor and analyze stakeholder' opinions (Pang and Lee, 2008). The affective orientation of a tweet is defined as its sentiment valence score and is measured as the average of the word's valence in the tweet normalized with the length of the message (see Appendix 3).

Results

RQ1. *What are the characteristics of CSR communication in social media?*

With the purpose of exploring the characteristics of CSR communication in social media, I first studied the CSR online community. The goal was to assess whether there is a well-connected CSR community where companies share the public, or if the overall community is less connected and companies act like hubs, connecting their specific audiences. In order to do so, I analyzed the percentage of nodes (i.e. users) shared by the companies. The investigation has shown a very low level of intersection between companies, implying the existence of several independent corporate-based networks rather than a unique well-connected CSR network. The highest level of shared nodes is 11.45 percent between Microsoft and Campbell, while the worst connected companies

are Green Mountain Coffee Roasters and Microsoft, which do not share any users. On average, companies shared 4 percent of their audiences. This finding suggests that companies do not share a common public; rather, they tend to develop their own audiences. The same result holds for the specific audiences, which tend to create separated networks, suggesting an interest in a specific company rather than generic CSR interest.

Secondly, I investigated the communication strategies adopted by the corporations by analyzing their relational ties and communication flow established with stakeholders and experts. Surprisingly, no corporations are interacting with the experts. Therefore they are not adopting a mediated strategy. Table I shows that the corporations are clearly divided into two main categories:

(1) self-centered; and

(2) dialogical.

The self-centered companies do not reciprocate friendships and are embedded in a low cohesive network. They exhibit low degrees of interaction and stakeholder content diffusion within their networks. On average, only 6 percent of their tweets directly addressed users, and 15 percent diffused the audience's content. According to the theoretical model, this reflects the self-centered strategy; companies use online communication space in a hierarchical way for self-presentation purposes and to avoid direct engagement. An exception is represented by Ford, which shows a low level of interaction with its audience, although a high degree of reciprocity. The dialogical strategy represents those companies who have built a cohesive network and interact profusely with their stakeholders. Starbucks particularly shows the highest level of engagement with 90 percent of reciprocity and 49.17 percent of direct interactions. Since there was no evidence of the mediated strategy in the analysis I only considered the self-centered and the dialogical strategies.

RQ2. Is there an alignment between companies CSR agenda and stakeholders social expectations in social media? If there is alignment, do the different communication strategies lead to different alignment outcomes?
Once the different strategies were identified, I applied topic analysis in order to investigate the discussions between stakeholders and companies and to evaluate whether they shared themes and vocabulary, and whether there are relevant differences between the two strategies. Table II shows the topics identified with the relative terms and the generic themes associated.

Company name	Degree of reciprocity (percent)	Degree of conversation (percent)	Communication strategy
Ford	94.63	16.86	Self-centered
Starbucks	89.25	49.17	Dialogical
Smarter Planet	70.24	53.61	Dialogical
Green Mountain Coffee Roasters	45.15	56.43	Dialogical
Xcel Energy	33.61	27.41	Self-centered
Campbell	14.50	11.90	Self-centered
Microsoft	1.90	33.45	Self-centered

Table I.
Communication
strategies by companies
based on relational and
conversational features

238

Table II.
Topics and keywords
detected

Theme	Topics	Keywords
Information seeking	Energy & Green & Us	Power, wind and solar associated with free, look, share
Information seeking	Thank & Share	Energy and solar with thank, share, great, day
Sustainable debate	Energy & Solar & Renew	Green, solar, energy but both without associations with news or business
Sustainable debate	Health & Care	Us, health, energy, solar, care
Market	Free & Business & Make	Business, make, market, work, blog
Democracy	Democracy & Support	Twitter, support, show, Iran, democracy

The first two topics are related to information seeking. The most important keywords are "energy", "power", "wind" and "solar" associated with "check", "project", "share", "see", "save". Topics 3 and 4 capture mainly the sustainability debate. Indeed, one topic refers to green and solar energy and the other to health and care, but both without any associations to news or business. In contrast, in topic 5, the most important words are "business", "make", "market", "work" and "use", indicating an interest in the market potential of green and alternative business. The last topic, with keywords like "democracy" and "support" associated with "Iran-", captures the large social support that Twitter users have shown to the Iranian protests in 2009, the so-called "Twitter Revolution". It is very interesting to notice how the focus of the CSR community is not only around environmental and health aspects, but also to political and ethical discussions within society as a whole. However, since topic analysis describes an association of words belonging to the same concept, it is interesting to examine if the vocabulary used by stakeholders and companies reflects different features of the concepts. Figure 2 represents the tag-clouds by groups, i.e. the set of most frequent words used by companies and audiences, in the self-centered and dialogical strategies, respectively.
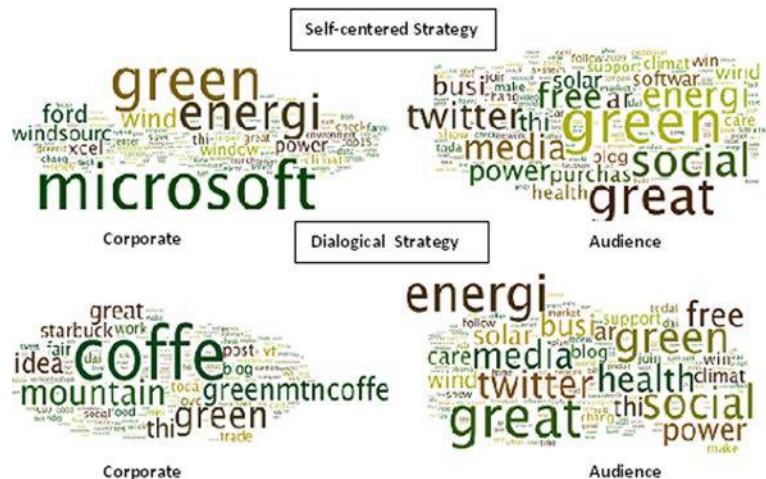

Figure 2.
Self-centred and dialogical strategy: most frequent words for companies and audience

Figure 2 shows very similar patterns among companies, regardless of the types of communication followed. Indeed, the two companies' clouds are composed of a few important words and a plethora of small words, while the audiences' clouds are more balanced and several words are important components of the discourse. The most frequent words in the companies' clouds are their names associated with CSR keywords, such as "green", "energy", "social", "power" and "work"; while the audience's tags range from "great", "energy", "health", "business", "care", "climate", "change" to "purchase", "business" and "free". In other words, while the concepts and related vocabulary do not vary strongly among groups, the main difference is that companies associate their brands to these concepts.

Next, I investigated organizational legitimacy as the congruence between CSR agenda and social expectations by measuring attention and affective orientation towards topics by group and by communication strategies. Figures 3 and 4 show the
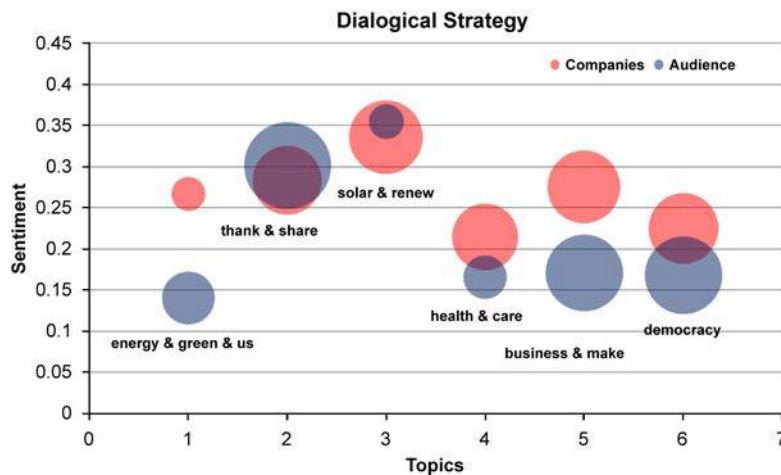

Figure 3.
Dialogical strategy:
attention and affective
orientation's congruence
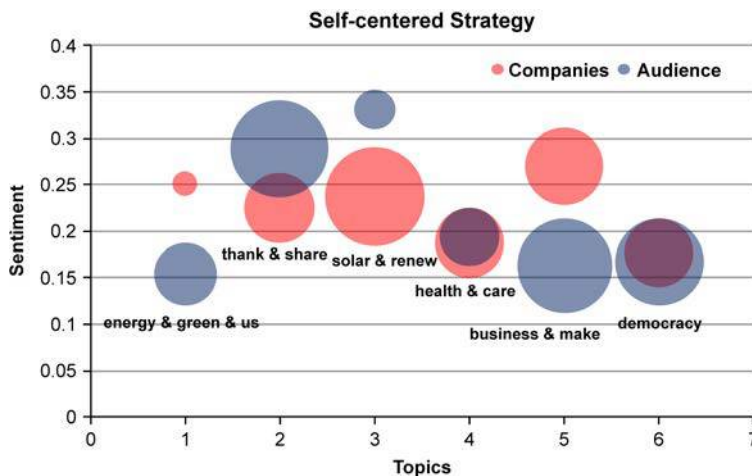by topic, audience and
companies

Figure 4.
Self-centred strategy:
attention and affective
orientations congruence
by topic, audience and
companies

different alignment outputs between companies and their audiences. The vertical axis represents the affective orientation about a topic, the horizontal axis represents the topics, and the size of the bubbles represents the level of attention towards a topic. First, I noticed that the sentiment is always positive, for both audiences and companies. This is coherent with other studies of online communities, which have shown how people tend to associate themselves with positive content (see Hansen *et al.*, 2011; Berger and Milkman, 2012). However, the variation among topics and groups is very limited, suggesting the need for further research into whether sentiment can be more informative if positive and negative opinions are considered separately. Figures 3 and 4 show that the information seeking process is important for the audience, but very marginal for the companies, in both strategies.

Topics 3 and 4 captured most of the companies' attention and a marginal amount of the audiences. Business related discussions drive almost 25 percent of total audience discourse, and almost 20 percent of the companies' attention; similarly, the theme of democracy in Iran accounts for 22 percent of audiences' and 17 percent of the companies' conversations.

Comparing the two strategies in terms of the ability to lead to alignment, it is evident that neither the information nor the dialogical strategy is able to establish a real convergence. However, the dialogical strategy shows more balance towards the companies' topics attention. This probably reflects the interactive nature of this approach that lead to conversation more orientated to the interests of both the companies and the audiences.

Discussion

The aim of this paper was to generate insights into the corporate communication field by applying new data-mining algorithms to investigate organizational legitimacy. Legitimacy was defined as the congruence of corporate CSR agenda and stakeholders' social expectations as emerging from different communication strategies in social media. The findings are two-fold and concern the discursive practices of the companies and the performances of the different communication strategies.

First, I have analyzed the discursive practices in terms of relevance of topics and vocabulary used by the companies and their audiences. I found that the most relevant topics for the audiences are related to information seeking. Information seeking is one of the most common patterns of usage of social media sites based on peer-to-peer information sharing (Java *et al.*, 2007). Stakeholders ask their peers about information and opinions on specific topics related to CSR, such as green energy and solar power. In contrast, the companies show a very "unnatural" pattern of usage. They are very keen on discussing the general CSR issues and are very marginal in the information seeking process, regardless of the communication strategy used. These findings point in the direction of what Schultz and Wehmeier (2010) call a "moralization" of corporate communications: "Corporations developed idealistic definitions of CSR within their self-presentations and refer to ethical or civic accounts and general social or environmental ills, such as climate change" (Schultz and Wehmeier, 2010, p. 18). The same result holds when analyzing the vocabulary: the messages are broadcast mostly with a direct mention of the company's name associated with positive CSR content, which is reminiscent of traditional advertising strategies. Secondly, I have examined the ability of the different communication strategies to align to stakeholders' social

expectations and increase corporate legitimacy. Following Habermas's assumption of deliberative public sphere, several authors seem to generally associate engagement and interaction with the creation of a dialogue and the co–creation of values and meanings, and therefore suggest that corporations must engage directly with stakeholders in order to increase legitimacy (Kent and Taylor, 1998; Moreno and Capriotti, 2009; Fieseler *et al.*, 2009). However, as pointed out by Schultz and Wehmeier (2010), these approaches "are based on the rationalist paradigm that argues both in a linear and causal manner" (Schultz and Wehmeier, 2010, p. 20). A simple information strategy can easily become a mere symbolic communication; in contrast, a more intensive claim can increase skepticism and attract critical attention (Ashforth and Gibbs, 1990). Furthermore, the dialogical strategy is based "on rational argumentations in the Habermasian way, but [social reality] is often constructed by emotional and moral communications" (Schultz and Wehmeier, 2010, p. 21). Schultz and Wehmeier (2010) have pointed out how "such emotionalized communication is often used by protest movements and does not aim at finding consensus, but at breaking up dialogue and, therefore, the mutual negotiation of social reality" (Schultz and Wehmeier, 2010, p. 21). As suggested by Lyotard (1984), the public sphere is often based on several publics, reflecting social inequalities, which organize themselves to reach public attention and dialogue is based on anarchy, individuality, disagreement and heated conversations. Investigating online conversations in social media sites, Papacharissi (2004) has found results consistent with Lyotard's definition of the public sphere. Assuming a non-a priori communicational assumption, I found evidence that supports this latter view. The results clearly show that establishing an interactive and engaging approach is not enough to align to stakeholders' social expectations. The idea that the more the dialog, the more the communality seems to fail to portray the complexity of the communication dynamics, such as the persistence of different vocabulary within the same discourse, or simply a dialog without alignment, as suggested by Lyotard (1984). At first glance, the lack of alignment seems to be related to corporations' communication practices, which follow traditional advertising strategies; however, further research must be carried out. Particularly, a more in-depth analysis of the communication practices can shed light on two main issues arising from this study. Firstly, if the corporate communicational format emerging from the empirical investigation of the CSR community (i.e. tendency to share generic CSR positive content associated with the company's name) represent the general corporate online presence, or if this finding only relates to the communication of CSR content. Secondly, empirical research needs to be carried out to investigate what drives alignment. Additionally, considering positive and negative sentiments separately might lead to new insight, particularly when analyzing crisis communication, where it is very likely to have two conflicting parties developing contrasting arguments.

Overall, this study has highlighted the great contribution that new data mining techniques, such as sentiment and topic analysis, can provide to the corporate communication field and to companies as well, particularly when applied to Internet communication. Indeed, these techniques allow the analysis of large amounts of data, which it would not be possible to investigate otherwise. Most importantly, sentiment analysis allows the automated track of online opinions and affect towards topics as well as companies. Likewise, latent semantic analysis allows investigation of the relevant topics as perceived by the stakeholders. Furthermore, new algorithms are emerging that increasingly exploit the informative content of online communication.

For instance, anomaly detection algorithms, traditionally applied for web-security or medical purposes, are increasingly applied to monitor the emergence of discourse "anomalies" on Twitter (see Sakaki *et al.*, 2010). These algorithms can be applied to monitor a corporation's opinion trends, predicting the emergence of potential scandals.

There are several ways that companies can foster alignment with their stakeholders on Twitter. As the present research shows, companies can use publicly available social media data to understand the conversations perceived as relevant and the opinions about specific topics of their audiences and subsequently design proper campaigns. For instance, the analysis has shown how users often seek CSR-related information, take solar panel and green habits for example; a demand that is not captured by the companies, regardless of the communication strategies adopted. It emerges clearly from this study the complexity of stakeholders' view and their high ethical expectations towards CSR, which a company should be aware of in order to reach a real alignment.

References

Andriof, J. and Waddock, S. (2002), "Unfolding stakeholder engagement", in Andriof, J., Waddock, S., Husted, B. and Rahman, S.S. (Eds), *Unfolding Stakeholder Thinking: Theory, Responsibility and Engagement*, Greenleaf Publishing, Sheffield, pp. 19-42.

Ashforth, B. and Gibbs, B. (1990), "The double-edge of organizational legitimation", *Organization Science*, Vol. 1 No. 2, pp. 177-94.

Basu, K. and Palazzo, G. (2008), "Corporate social responsibility: a process model of sensemaking", *Academy of Management Review*, Vol. 33 No. 1, pp. 122-36.

Beck, U. (2000), *Risk Society. Towards a New Modernity*, Sage Publications, Thousand Oaks, CA.

Berger, J. and Milkman, K. (2012), "What makes online content viral?", *Journal of Marketing Research*, Vol. 49 No. 2, pp. 192-205.

Berger, P.L. and Luckman, L. (1966), *The Social Construction of Reality: A Treatise on the Sociology of Knowledge*, Anchor, New York, NY.

Bhattacharya, C.B. and Sen, S. (2004), "Doing better by doing good: when, why, and how consumer respond to corporate social initiatives", *California Management Review*, Vol. 47 No. 1, pp. 9-24.

Borgatti, S. (2005), "Centrality and network flow", *Social Networks*, Vol. 27 No. 1, pp. 55-71.

Botan, C. (1997), "Ethics in strategic communication campaigns: the case for a new approach to public relations", *Journal of Business Communication*, Vol. 34 No. 2, pp. 188-202.

Boyd, D., Golder, S. and Lotan, G. (2010), "Tweet, tweet, retweet: conversational aspects of retweeting on Twitter", *Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS 43), Kauai, HI, January*, IEEE Computer Society, Piscataway, NJ, pp. 1-10.

Capriotti, P. and Moreno, A. (2007), "Corporate citizenship and public opinions: the importance and interactivity of social responsibility issues on corporate websites", *Public Relations Review*, Vol. 33 No. 1, pp. 84-91.

Carey, J.W. (1989), *Communication as Culture: Essay on Media and Society*, Routledge, New York, NY.

Castells, M. (1997), *The Power of Identity*, *The Information Age: Economy, Society and Culture*, Vol. II, Blackwell, Cambridge, MA.

Colleoni, E. (2012), "New forms of digital marketing", in Belk, R. and Liama, R. (Eds), *The Routledge Companion to Digital Consumption*, Routledge, New York, NY, pp. 124-34.

Colleoni, E., Arvidsson, A., Hansen, L.K. and Marchesini, M. (2011), "Measuring corporate reputation using sentiment analysis", paper presented at the 15th International Conference on Corporate Reputation, Brand, Identity and Competitiveness, New Orleans, LA, May 18-20.

Corporate Responsibility Magazine (2009), "Corporate Responsibility Magazine's 100 Best Corporate Citizens List", available at http://thecro.com/content/corporate-responsibility-magazine%E2%80%99s-%E2%80%9C100-best-corporate-citizens-list%E2%80%9D

Coupland, C. (2005), "Corporate social responsibility as argument on the web", *Journal of Business Ethics*, Vol. 62 No. 4, pp. 355-66.

D'Aunno, T. and Zuckerman, H.S. (1987), "A life-cycle model of organizational federations: the case of hospitals", *Academy of Management Review*, Vol. 12 No. 3, pp. 534-45.

Dawkins, J. (2004), "Corporate responsibility: the communication challenge", *Journal of Communication Management*, Vol. 9 No. 2, pp. 108-19.

Dellarocas, C. (2003), "The digitization of word of mouth: promise and challenges of online feedback mechanisms", *Management Science*, Vol. 49 No. 10, pp. 1407-24.

Dowling, J. and Pfeffer, J. (1975), "Organizational legitimacy: social values and organizational behavior", *The Pacific Sociological Review*, Vol. 18 No. 1, pp. pp122-36.

Fieseler, C., Fleck, M. and Meckel, M. (2009), "Corporate social responsibility in the blogosphere", *Journal of Business Ethics*, Vol. 91, pp. 599-614.

Freeman, E. (1984), *Strategic Management: A Stakeholder Approach*, Pitman, Marshfield, MA.

Habermas, J. (1989), *The Structural Transformation of the Public Sphere*, MIT Press, Cambridge, MA, (first published 1962).

Hansen, L.K., Arvidsson, A., Nielsen, F., Colleoni, E. and Etter, M. (2011), "Good friends, bad news – affect and virality in Twitter", *Proceedings of the International Workshop on Social Computing, Network, and Services SocialComNet'2011*, Springer Communications in Computer and Information Science, Vol. 185, Springer, Berlin, pp. 34-43.

Herrmann, K.K. (2004), "Corporate social responsibility and sustainable development: the European Union initiative as a case study", *Indiana Journal of Global Legal Studies*, Vol. 11 No. 2, pp. 204-32.

Java, A., Song, X., Finin, T. and Tseng, B. (2007), "Why we Twitter: understanding microblogging usage and communities", *WebKDD/SNA-KDD '07: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, New York, NY*, pp. 56-65.

Jenkins, H. (2006), *Convergence Culture: Where Old and New Media Collide*, New York University Press, New York, NY.

Kent, M.L. and Taylor, M. (1998), "Building dialogic relationships through the world wide web", *Public Relations Review*, Vol. 24 No. 3, pp. 321-34.

Ingenhoff, D. and Koelling, M.A. (2009), "The potential of web sites as a relationship building tool for charitable fundraising NPOs", *Public Relations Review*, Vol. 35, pp. 66-73.

Insch, A. (2008), "Online communication of corporate environmental citizenship: a study of New Zealand's electricity and gas retailers", *Journal of Marketing Communications*, Vol. 14 No. 2, pp. 139-53.

Latour, B. and Weibel, P. (2005), *Making Things Public*, MIT Press, Cambridge, MA.

Lyotard, J.F. (1984), *The Postmodern Condition*, (trans. by Massumi, B. *et al.*), University of Minnesota Press, Minneapolis, MN.

Marwick, A. and Boyd, D. (2011), "I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience", *New Media and Society*, Vol. 13 No. 1, pp. 114-33.

Massey, J.E. (2001), "Managing organizational legitimacy: communication strategies for organizations in crisis", *The Journal of Business Communication*, Vol. 38 No. 2, pp. 153-83.

Mølgaard, L. and Szewczyk, M. (2010), "Sentiment analysis using machine learning", unpublished results, Department of Informatics and Mathematical Modeling, Technical University of Denmark.

Moreno, A. and Capriotti, P. (2009), "Communicating CST: citizenship and sustainability on the Web", *Journal of Communication Management*, Vol. 13 No. 2, pp. 157-75.

Morsing, M. and Schultz, M. (2006), "Corporate social responsibility communication: stakeholder information, response and involvement strategies", *Business Ethics: A European Review*, Vol. 15 No. 4, pp. 323-38.

Morsing, M., Schultz, M. and Nielsen, K.U. (2008), "The 'Catch 22' of communicating CSR: findings from a Danish study", *Journal of Marketing Communications*, Vol. 14 No. 2, pp. 97-111.

Nielsen, F.Å. (2011), "A new ANEW: evaluation of a word list for sentiment analysis in microblogs", unpublished results, Department of Informatics and Mathematical Modeling, Technical University of Denmark, Kongens Lyngby.

Palazzo, G. and Scherer, G. (2006), "Corporate legitimacy as deliberation: a communicative framework", *Journal of Business Ethics*, Vol. 66 No. 1, pp. 71-88.

Pang, B. and Lee, L. (2008), "Opinion mining and sentiment analysis", *Foundations and Trends in Information Retrieval*, Vol. 2 No. 1, pp. 1-135.

Papacharissi, Z. (2004), "Democracy online: civility, politeness, and the democratic potential of online political discussion groups", *New Media and Society*, Vol. 6 No. 2, pp. 259-83.

Parsons, P. (1960), *Structure and Process in Modern Societies*, The Free Press, Glencoe, IL.

Rafaeli, S. and Sudweeks, F. (1997), "Networked interactivity", *Journal of Computer-Mediated Communication*, Vol. 2 No. 4.

Ravikant, N. and Rifkin, A. (2010), "Why Twitter is massively undervalued compared to Facebook", available at: http://techcrunch.com/2010/10/16/why-twitter-is-massively-undervalued-compared-to-facebook/.

Ryan, M. (2003), "Public relations and the web: organizational problems, gender, and institution type", *Public Relations Review*, Vol. 29 No. 3, pp. 335-49.

Sakaki, T., Okazaki, M. and Matsuo, Y. (2010), "Earthquake shakes Twitter users: real-time event detection by social sensors", *Proceedings of WWW 2010, April 26-30, Raleigh, NC*.

Scherer, A. and Palazzo, G. (2007), "Toward a political conception of corporate responsibility: business and society seen from a Habermasian perspective", *Academy of Management Review*, Vol. 32 No. 4, pp. 1096-120.

Schultz, F., Utz, S. and Goritz, A. (2011), "Is the medium the message? Perceptions of and reactions to crisis communication via Twitter, blogs and traditional media", *Public Relations Review*, Vol. 37 No. 1, pp. 20-7.

Schultz, F. and Wehmeier, S. (2010), "Institutionalization of corporate social responsibility within corporate communications: combining institutional, sensemaking and communication perspectives", *Corporate Communications: An International Journal*, Vol. 15 No. 1, pp. 9-29.

Schultz, T. (2000), "Mass media and the concept of interactivity: an exploratory study of online forums and reader email", *Media, Culture and Society*, Vol. 22 No. 2, pp. 205-21.

Smith, C. (2003), "Corporate social responsibility: whether or how?", *California Management Review*, Vol. 45 No. 4, pp. 52-76.

Suchman, M.C. (1995), "Managing legitimacy: strategic and institutional approaches", *Academy of Management Review*, Vol. 20 No. 3, pp. 571-610.

Thelwall, M., Buckley, K. and Paltoglou, G. (2011), "Sentiment in Twitter events", *Journal of American Society for Information Science and Technology*, Vol. 62 No. 2, pp. 406-18.

Wood, D. (1991), "Corporate social performance revisited", *Academy of Management Review*, Vol. 16 No. 4, pp. 691-718.

Zucker, L. (1987), "Institutional theories of organizations", *Annual Review of Sociology*, Vol. 13, pp. 443-644.

Further reading

Arvidsson, A. and Colleoni, E. (2012), "Value in informational capitalism and on the internet. A reply to Christian Fuchs", *The Information Society*, Vol. 28 No. 3, pp. 135-50.

Carroll, A.B. (1979), "A three-dimensional conceptual model of corporate performance", *Academy of Management Review*, Vol. 4 No. 4, pp. 497-505.

Esrock, S.L. and Leichty, G.B. (1998), "Social responsibility and corporate web pages: presentation or agenda setting?", *Public Relations Review*, Vol. 24 No. 3, pp. 305-19.

Gerlitz, C. and Helmond, A. (2011), "Hit, link and share: organizing the social and the fabric of the web in a like economy", paper presented at the DMI Mini-Conference, University of Amsterdam, Amsterdam, January 24-25.

Honeycutt, C. and Herring, S.C. (2009), "Beyond microblogging: conversation and collaboration via Twitter", *Proceedings of the 42nd Hawaii International Conference on System Sciences (HICSS 42)*, IEEE Computer Society, Washington, DC, pp. 1-10.

Podnar, K. (2008), "Guest editorial: communicating corporate social responsibility", *Journal of Marketing Communications*, Vol. 14 No. 2, pp. 75-81.

Porter, M. (2006), "Mapping social opportunities", *Harvard Business Review*, Vol. 84 No. 12, pp. 86-7.

Powell, W. and DiMaggio, P. (1991), *The New Institutionalism in Organizational Analysis*, University of Chicago Press, Chicago, IL.

Scott, R.W. (1987), *Organizations: Rational, Natural, and Open Systems*, Prentice Hall, Englewood Cliffs, NJ.

Weick, K. (1995), *Sensemaking in Organizations*, Sage Publications, Thousand Oaks, CA.

Appendix 1. Training the CSR classifier

In order to detect CSR-related tweets automatically for the whole network, we trained an automated classifier with a machine learning algorithm based on a database of manually classified CSR tweets. In order to do so, first, one of the authors, drawing on knowledge developed as a CSR and communication scholar, manually classified 3,000 tweets according to the distinction CSR content/non-CSR content. As a test of inter-coder reliability, the other author and a colleague of the department working with the same database respectively and independently coded sample sections. No significant differences were identified and the initial coding was found to be internally consistent. From the pre-classified database we extracted the 6,457 most frequent words and used those words as terms. We used this information as a base to "train" a data-mining algorithm to classify the tweets automatically. We used the so-called naive Bayes classifier, where the probability of a tweet being CSR-related is expressed as the combined probabilities of each word to be CSR related in the tweet.

We split the corpus of data in 75 percent for training and the rest for testing and we reached an average accuracy of 89.33 percent, i.e. percentage of correct classification of CSR/non-CSR related tweets. Particularly, we classified correctly CSR content in 89 cases out of 100 and non-CSR content in 92 cases out of 100.

Appendix 2. Latent semantic analysis

Latent semantic analysis is a natural language processing technique based on a singular value decomposition method that identifies relationships between set of documents by identifying their shared concepts. LSA maps all the documents and the terms within them as a whole into a "concept" space and compares them in this space. These algorithms assume that words that are close in meaning will occur close together in text. For instance, using a search engine search for "sand", documents are returned that do not contain the search term "sand" but contain terms like "beach". Therefore, the model trained on such a corpus has identified a latent relation: sand is semantically close to beach. These terms share the concept space and will be therefore classified under the same topic. One can view LSA as a transformation from one coordinate system (documents × terms) to another – concept space with reduced dimensions.

The document term matrix X is decomposed into matrix Σ containing singular values and left (U) and right ($_v$) singular vectors. By limiting number of dimensions (by setting singular values in Σ to 0) one can create a "concept" space. This concept space is related to both documents and terms. Therefore, one can see, for example, what terms form a concept, which can be viewed as topics. The LSA model was trained on corpus of 326,291 documents. It contained 43,151 unique terms. In total there were 2,815,282 non-zero entries. We have identified seven dimensions, referring to six different topics and one residual component (see Table AI). The explanation power strongly decreases over topics, which means that the topics well discriminate the information.

Appendix 3. Sentiment analysis

Sentiment analysis allows subjective perceptions to acquire an objective existence as observable and measurable forms (Colleoni, 2012).

The most common approach for affective text classification is using a lexicon. This approach requires "the creation of a knowledge base-lexicon of affective words, with additional data characterizing emotional states and relations" (Mølgaard and Szewczyk, 2010, p. 2). The most general sentiment classification allows the polarity classification of a text by distinguishing between positive and negative sentiment. More elaborated classifications include the identification of the strength of a sentiment. The underlying assumption is that "humans can differentiate between mild and strong emotions" (Thelwall *et al.*, 2011, p. 4). Sentiment expressions are classified according to their valence – i.e. how positive or negative the expressed sentiment is – and arousal – i.e. level of the emotional excitement (Colleoni *et al.*, 2011). In this work, we used a word list developed by Nielsen (2011) based on Twitter text that includes internet slang and obscene words. This Twitter-based word list has 2,477 unique words and includes 15 phrases; the score ranges from -5 (very negative) to $+5$ (very positive). For instance, the word "adorable" is classified as a positive term with valence 3, while "aggression" is negative with valence -2. The valence of a tweet is computed as the sum of the valences of the individual words in the tweet, while the arousal the sum of the absolute value of the valences.

1. Energy & Green & Us	Topic #1 (143.076): −0.684*"social" + −0.533*"media" + 0.218*"energi" + 0.216*"great" + 0.158*"green" + 0.118*"us" + 0.109*"solar" + 0.095*"power" + 0.088*"wind" + 0.073*"thank" + 0.070*"make" + 0.069*"day" + −0.053*"market" + 0.051*"free" + 0.048*"go" + −0.046*"network" + 0.043*"one" + 0.041*"save" + 0.041*"health" + 0.039*"check" + 0.038*"climat" + 0.035*"look" + 0.034*"home" + 0.033*"renew" + 0.031*"time" + 0.030*vwork" + 0.030*"care" + 0.029*"car" + 0.029*"win" + 0.029*"see"
2. Thank & Share	Topic #2 (133.750): 0.681*"great" + −0.488*"energi" + −0.252*"solar" + −0.240*"power" + −0.225*"green" + −0.201*"wind" + 0.142*"thank" + −0.073*"renew" + −0.067*"social" + −0.053*"media" + 0.047*"check" + −0.042*"save" + −0.042*"home" + −0.040*"plan" + −0.040*"climat" + −0.039*"clean" + −0.038*"busi" + −0.035*"build" + −0.034*"go" + −0.031*"news" + 0.030*"articl" + −0.030*"project" + −0.029*"compani" + −0.029*"make" + 0.029*"day" + −0.028*"#green" + −0.027*"bill" + −0.027*"chang" + 0.026*"share" + −0.026*"turbin"
3. Energy & Solar & Renew	Topic #3 (121.217): 0.532*"great" + 0.449*"energi" + −0.282*"us" + 0.235*"solar" + 0.207*"power" + 0.185*"wind" + −0.177*"free" + −0.168*"health" + 0.119*"social" + −0.117*"make" + −0.116*"care" + −0.115*"twitter" + −0.098*"join" + −0.097*"win" + −0.097*"day" + 0.096*"media" + −0.084*"busi" + −0.080*"one" + −0.070*"chang" + 0.066*"renew" + −0.062*"u" + −0.062*"follow" + −0.059*"pleas" + −0.059*"want" + −0.058*"good" + −0.058*"blog" + −0.057*"climat" + −0.057*"check" + −0.055*"go" + −0.055*"need"
4. Us & Health & Care	Topic#4(104.873) : −0.846*"green" + 0.312*"us" + 0.155*"energi" + 0.138*"power" + 0.136*"health" + −0.121*"go" + 0.116*"wind" + 0.115*"solar" + −0.104*"busi" + 0.090*"care" + −0.090*"blog" + 0.079*"join" + −0.065*"job" + 0.060*"free" + 0.057*"thank" + 0.052*"climat" + 0.051*"chang" + −0.042*"great" + 0.041*"twitter" + −0.041*"#green" + −0.036*"build" + 0.033*"share" + 0.033*"make" + 0.031*"u" + −0.028*"way" + −0.028*"product" + 0.026*"follow" + 0.025*"pleas" + 0.024*"renew" + 0.023*"bill"
5. Free & Business & Make	Topic #5 (85.177): −0.610*"us" + −0.296*"green" + −0.272*"health" + 0.252*"free" + 0.251*"twitter" + 0.230*"busi" + −0.185*"care" + 0.180*"make" + 0.157*"blog" + 0.113*"day" + −0.107*"join" + −0.106*"great" + 0.097*"one" + −0.096*"social" + −0.093*"climat" + 0.079*"market" + −0.078*"thank" + 0.076*"power" + −0.075*"media" + 0.069*"use" + 0.068*"good" + 0.066*"best" + 0.064*"work" + 0.059*"onlin" + 0.054*"win" + 0.053*"home" + 0.052*"solar" + 0.051*"way" + −0.048*"energi" + 0.048*"offer"

(*continued*)

Table AI.
Topic extraction and
relevant words

Table AI.

6. Democracy & Support	Topic #6 (62.198): 0.700*"twitter" + −0.275*"health" + 0.220*"us" + 0.192*"thank" + 0.182*"support" + −0.181*"care" + 0.154*"show" + 0.138*"iran" + 0.133*"add" + −0.125*"busi" + 0.118*"avatar" + 0.115*"democraci" + −0.114*"climat" + −0.114*"blog" + −0.107*"make" + 0.103*"overlay" + 0.091*"follow" + 0.089*"green" + −0.089*"chang" + −0.079*"market" + −0.076*"day" + 0.074*"join" + −0.060*"good" + −0.058*"need" + −0.057*"one" + −0.057*"work" + −0.054*"go" + 0.054*"energi" + 0.053*"pratyush" + 0.052*"post"
7. Residual component	Topic #0 (168.005): −0.352*"great" + −0.329*"social" + −0.252*"media" + −0.183*"us" + −0.169*"green" + −0.166*"energi" + −0.151*"make" + −0.141*"thank" + −0.137*"busi" + 0.127*"blog" + −0.121*"twitter" + −0.120*"free" + −0.116*"good" + −0.115*"day" + −0.112*"power" + −0.111*"market" + −0.107*"check" + −0.107*"look" + −0.107*"one" + −0.097*"go" + −0.095*"work" + −0.091*"need" + −0.090*"way" + −0.089*"solar" + −0.083*"time" + −0.081*"see" + −0.079*"use" + −0.078*"chang" + −0.076*"join" + −0.074*"peopl"

About the authors

Elanor Colleoni is a Postdoctoral Researcher at the Department of Intercultural Communication and Management, Copenhagen Business School. With a background in computer science and sociology, she is currently working on the project "Responsible Business in the Blogosphere" (RBB). She is investigating the relationship between information diffusion and emotional content in social media and the impact of social relations on brand and corporate reputation in new media. Elanor Colleoni can be contacted at: elc.ikl@cbs.dk

This article has been cited by:

1. Yuli Suárez-Rico, Mauricio Gómez-Villegas, María García-Benau. 2018. Exploring Twitter for CSR Disclosure: Influence of CEO and Firm Characteristics in Latin American Companies. *Sustainability* **10**:8, 2617. [Crossref]

2. Jing Yang, Kelly Basile, Olivia Letourneau. 2018. The impact of social media platform selection on effectively communicating about corporate social responsibility. *Journal of Marketing Communications* **49**, 1-23. [Crossref]

3. RambocasMeena, Meena Rambocas, PachecoBarney G., Barney G. Pacheco. 2018. Online sentiment analysis in marketing research: a review. *Journal of Research in Interactive Marketing* **12**:2, 146-163. [Abstract] [Full Text] [PDF]

4. Swaleha Peeroo, Martin Samy, Brian Jones. 2018. Trialogue on Facebook pages of grocery stores: Customer engagement or customer enragement?. *Journal of Marketing Communications* **2**, 1-23. [Crossref]

5. AraujoTheo, Theo Araujo, KollatJana, Jana Kollat. 2018. Communicating effectively about CSR on Twitter. *Internet Research* **28**:2, 419-431. [Abstract] [Full Text] [PDF]

6. Francisca Farache, Isobel Tetchner, Jana Kollat. 145. [Crossref]

7. Constance E. Kampf. 2018. Connecting Corporate and Consumer Social Responsibility Through Social Media Activism. *Social Media + Society* **4**:1, 205630511774635. [Crossref]

8. Aimei Yang, Wenlin Liu. 2018. CSR Communication and Environmental Issue Networks in Virtual Space: A Cross-National Study. *Business & Society* 000765031876356. [Crossref]

9. Michael Etter, Elanor Colleoni, Laura Illia, Katia Meggiorin, Antonino D'Eugenio. 2018. Measuring Organizational Legitimacy in Social Media: Assessing Citizens' Judgments With Sentiment Analysis. *Business & Society* **57**:1, 60-97. [Crossref]

10. Subhajit Bhattacharya. 2017. Does corporate social responsibility contribute to strengthen brand equity? An empirical study. *International Review on Public and Nonprofit Marketing* **14**:4, 513-533. [Crossref]

11. Ebru Uzunoğlu, Selin Türkel, Burcu Yaman Akyar. 2017. Engaging consumers through corporate social responsibility messages on social media: An experimental study. *Public Relations Review* **43**:5, 989-997. [Crossref]

12. Mark Anthony Camilleri. 2017. The Promotion of Responsible Tourism Management Through Digital Media. *Tourism Planning & Development* **2009**, 1-19. [Crossref]

13. KollatJana, Jana Kollat, FaracheFrancisca, Francisca Farache. 2017. Achieving consumer trust on Twitter via CSR communication. *Journal of Consumer Marketing* **34**:6, 505-514. [Abstract] [Full Text] [PDF]

14. Xiaorui Wang. 2017. Regulatory compliance as fulfilment of corporate social responsibility: an interpretative textual analysis on sustainability reports of two Chinese listed agribusinesses. *Asian Journal of Sustainability and Social Responsibility* **2**:1, 23-40. [Crossref]

15. WiencierzChristian, Christian Wiencierz, RöttgerUlrike, Ulrike Röttger. 2017. The use of big data in corporate communication. *Corporate Communications: An International Journal* **22**:3, 258-272. [Abstract] [Full Text] [PDF]

16. Lina M. Gómez, Ramón W. Borges-Tavárez. CSR Online Communication in Latin America: An Analysis of Social Media Platforms 113-132. [Abstract] [Full Text] [PDF] [PDF]

17. Moonhee Cho, Lauren D. Furey, Tiffany Mohr. 2017. Communicating Corporate Social Responsibility on Social Media. *Business and Professional Communication Quarterly* **80**:1, 52-69. [Crossref]

18. Doula Zaharopoulos, Linchi Kwok. 2017. Law Firms' Organizational Impression Management Strategies on Twitter. *Journal of Creative Communications* **12**:1, 48-61. [Crossref]

19. Merkl-DaviesDoris M., Doris M. Merkl-Davies, BrennanNiamh M., Niamh M. Brennan. 2017. A theoretical framework of external accounting communication. *Accounting, Auditing & Accountability Journal* **30**:2, 433-469. [Abstract] [Full Text] [PDF]

20. Gregory D. Saxton, Lina Gomez, Zed Ngoh, Yi-Pin Lin, Sarah Dietrich. 2017. Do CSR Messages Resonate? Examining Public Reactions to Firms' CSR Efforts on Social Media. *Journal of Business Ethics* . [Crossref]

21. OlkkonenLaura, Laura Olkkonen. 2017. A conceptual foundation for expectations of corporate responsibility. *Corporate Communications: An International Journal* **22**:1, 19-35. [Abstract] [Full Text] [PDF]

22. Andreas Deckmann, Franziska Freudenberger, Silke Bustamante, Andrea Pelzeter, Rudi Ehlscheidt. Dienstleistungssektor: Mitarbeiterbindung und CSR-Kommunikation 293-306. [Crossref]

23. Urša Golob, Nataša Verk, Klement Podnar. Knowledge Integration in the European CSR Communication Field: An Institutional Perspective 273-291. [Crossref]

24. Mark Anthony Camilleri. Unlocking Corporate Social Responsibility Through Integrated Marketing Communication 41-59. [Crossref]

25. Lucinda L. Austin, Barbara Miller Gaither. 2016. Examining Public Response to Corporate Social Initiative Types. *Social Marketing Quarterly* **22**:4, 290-306. [Crossref]

26. Barbara Miller Gaither, Lucinda Austin. 2016. Campaign and corporate goals in conflict: Exploring company-issue congruence through a content analysis of Coca-Cola's twitter feed. *Public Relations Review* **42**:4, 698-709. [Crossref]

27. Bahadır Birim. 2016. Evaluation of Corporate Social Responsibility and Social Media as Key Source of Strategic Communication. *Procedia - Social and Behavioral Sciences* **235**, 70-75. [Crossref]

28. TamperePäivi, Päivi Tampere, TampereKaja, Kaja Tampere, Luoma-AhoVilma, Vilma Luoma-Aho. 2016. Facebook discussion of a crisis: authority communication and its relationship to citizens. *Corporate Communications: An International Journal* **21**:4, 414-434. [Abstract] [Full Text] [PDF]

29. John Brummette, Lynn M. Zoch. 2016. How stakeholders' personal values influence their value expectations for legitimate organizations. *Corporate Communications: An International Journal* **21**:3, 309-321. [Abstract] [Full Text] [PDF]

30. Pantelitsa P. Eteokleous, Leonidas C. Leonidou, Constantine S. Katsikeas. 2016. Corporate social responsibility in international marketing: review, assessment, and future research. *International Marketing Review* **33**:4, 580-624. [Abstract] [Full Text] [PDF]

31. Jari Lyytimäki, Lasse Peltonen. 2016. Mining through controversies: Public perceptions and the legitimacy of a planned gold mine near a tourist destination. *Land Use Policy* **54**, 479-486. [Crossref]

32. FarivarFarveh, Farveh Farivar, Scott-LaddBrenda, Brenda Scott-Ladd. 2016. Growing corporate social responsibility communication through online social networking in Iran. *International Journal of Organizational Analysis* **24**:2, 274-290. [Abstract] [Full Text] [PDF]

33. Ganga Sasidharan Dhanesh. 2015. Corporate social responsibility (CSR) in India: A dialectical analysis of the communicative construction of the meanings and boundaries of CSR in India. *Public Relations Inquiry* **4**:3, 287-303. [Crossref]

34. Mark Kilgour, Sheila L. Sasser, Roy Larke. 2015. The social media transformation process: curating content into strategy. *Corporate Communications: An International Journal* **20**:3, 326-343. [Abstract] [Full Text] [PDF]

35. Paula Maria Bögel. 2015. Processing of CSR communication: insights from the ELM. *Corporate Communications: An International Journal* **20**:2, 128-143. [Abstract] [Full Text] [PDF]

36. Imran Ali, Ana Isabel Jiménez-Zarco, Marta Bicho. Using Social Media for CSR Communication and Engaging Stakeholders 165-185. [Abstract] [Full Text] [PDF] [PDF]

37. Shuo Yao, John Brummette, Yi Luo. 2015. Balancing between organizations and society: exploring Chinese organizations' legitimacy efforts. *Corporate Communications: An International Journal* **20**:1, 90-110. [Abstract] [Full Text] [PDF]

38. Andrea Pérez. 2015. Corporate reputation and CSR reporting to stakeholders. *Corporate Communications: An International Journal* **20**:1, 11-29. [Abstract] [Full Text] [PDF]

39. Michael Czinkota, Hans Ruediger Kaufmann, Gianpaolo Basile. 2014. The relationship between legitimacy, reputation, sustainability and branding for companies and their supply chains. *Industrial Marketing Management* **43**:1, 91-101. [Crossref]

40. Maria Besiou, Mark Lee Hunter, Luk N. Van Wassenhove. 2013. A Web of Watchdogs: Stakeholder Media Networks and Agenda-Setting in Response to Corporate Initiatives. *Journal of Business Ethics* **118**:4, 709-729. [Crossref]

41. Itziar Castelló, Mette Morsing, Friederike Schultz. 2013. Communicative Dynamics and the Polyphony of Corporate Social Responsibility in the Network Society. *Journal of Business Ethics* **118**:4, 683-694. [Crossref]

42. André Grützmann, Cleber Carvalho de Castro, Anderson Antonio Freire de Moraes Meireles, Renan Carlos Rodrigues. Organizational Architecture and Online Social Networks 508-524. [Crossref]

43. Melis Kaytaz Yiğit, Özge Kirezli. Understanding Historical Background of Corporate Social Responsibility (CSR) and Realizing Social Media as a New Horizon in CSR Communication 59-85. [Crossref]