Social media and sentiment in bioenergy consultation

Introduction

Growth within the bioenergy and EfW (energy from waste) sectors within the UK is representative of European and worldwide trends. There is an ever-increasing need to divert organic waste from landfill in light of the European landfilling directive (1999/31/EC) and the rapidly decreasing landfilling capacity bank. A recent review of the incinerators within the UK found that of the 25 reviewed municipal solid waste (MSW) EfWs that there is a conservative approach in selecting the best available technology and efficient use of the produced energy (Nixon, Wright et al. 2013). Furthermore, most of the large scale EfW facilities of this type are financed in partnership with Local Authorities under Private Finance Initiative (PFI) or Public Private Partnerships (PPP) contracts over a 20 to 25 year term. EfW development is a necessity for Local Authorities as a method for handling the large amount of MSW produced, however, there is a general disconnect with the local population over the development of waste management solutions. The public generally have a negative association with the waste management sector that is normally circumvented with an ‘out of sight, out of mind’ mentality. This is not possible when there is a proposed development of an EfW facility is in close proximity to any populated places. Bioenergy developments that handle only the organic fraction of residue biomass or virgin material often suffer with the same labelling as waste management and incineration.

Bioenergy and energy from waste (EfW) projects frequently fail during the project development stage due to objections from local residents and activist groups. Objections against bioenergy and energy from waste can be broadly split between local and global objections depending on the nature of the objection. Local opposition is common for most classifications of private development in the UK but bioenergy and EfW projects have their
own particular blend of opposition as quantified in detail by (Upham and Shackley 2007).

Pollution, traffic, noise, odour and visual impact are all operation, technology and location specific concerns for local residents. Meanwhile there is a group of global ‘challenges’ for bioenergy and EfW projects to overcome as identified by many authors including fuel sustainability, legal and legislative requirements (for pollution control), rural job creation, habitat destruction and change, and issues around the food for fuel debate.

**Project Development Stages**

The development of energy projects within the UK follows a fairly standard process irrespective of whether they are biomass, waste or other renewables such as wind or solar.

The most significant difference for these is the scale of facility and whether the development requires waste handling and use. Biomass and waste schemes typically have an end to end development cycle of approximately 3 to 5 years. The more bespoke the scheme and at greater scales, with the handling of waste and hazardous materials, the longer the development phase and the greater the development costs.

The Department for Energy and Climate Change identifies the development phases as: prescoping, scoping, application, consent, logistics, development, operate and end of life (DECC 2011). During project development, it is necessary to consult with the local community as part of the planning process. Stakeholder views need to be addressed in order to establish democratic legitimacy for, and to build acceptance of, planning decisions with significant impacts. This is included as part of the ‘planning application and Environmental Impact Assessment’ activity in the application phase of development. It is during this activity that the developer must respond to the concerns of stakeholder, in particular the local community.
This is typically done using traditional communication tools, through events at the village hall, for example, if the development is within a rural area. The local community may also raise their concerns formally to the developer and/or planning authority either by a written letter or an email. Questionnaires (Upham & Shackley 2007) and combined questionnaire and focus groups (Upham, Shackley & Waterman 2007) have also been used where a more systematic analysis of local opinion is desired. The consultation process does not currently extend to informal concerns or sentiment (subjective information such as opinions, attitudes, and feelings expressed in text) through more recent communication channels such as social media. However, as there is an increasing internet presence for anti-development groups with local communities, it is likely that they will have a social media presence and therefore not only express their sentiment to energy developments but may also be influenced by others.

**Information Flow and Social Media**

Public consultation is not a straightforward process, and if mismanaged can backfire. The forms that consultation takes have been theorised as a “ladder of citizen participation” (Arnstein 1969), and as “information flows” (Stringer et al. 2006)(e.g. Barreteau et al. 2010). The rationale underlying these models is that participation should be maximized, and that this is achieved by creating flows of information back from citizens to policy makers and stakeholder organizations. For example, Hurlbert reports that early consultations on nuclear policy in Saskatchewan were selective, involving only interested stakeholders. Hence the resulting report was not well received by the public at large (Hurlbert 2014). Later consultations, involving much wider citizen engagement, which established communication flows among all the interested parties, were more successful.
‘Social Media’ can be defined as Web technologies which combine user generated content with social networks of ‘friends’ or ‘followers’ to target the content and get responses. Prominent examples are blogs, podcasts, microposting sites, of which Twitter has become the best known, forums, and social networking sites such as Facebook. These are having profound affects on the ways public communication is conducted, with journalism (Fahey & Nisbet 2011), marketing (Ellis-Chadwick 2009), and politics (Parmalee 2014) (Tumasjan et al. 2010) all adopting, and adapting to, the new technologies. Social media are also proposed to become part of the mix of consultation and participation tools for developing energy strategy, e.g., (Hurlbert 2014), because they are seen as having the potential to enable citizen participation: the ability to comment on blogs, retweet posts on Twitter, etc. facilitates dialogue among the parties.

Key benefits of social media to its users are: inclusiveness (social media tools are usually free to use and anyone can sign up for an account), low barriers to participation (it is easy to post a comment, ‘like’ a post, or tweet, even if authoring a blog is more of a commitment), and knowledge of the audience (provided by friends and followers lists). However, social media pose challenges for both citizens and organizations. As Boyd observes “when politicians and activists talk about using MySpace and Facebook, they aren’t talking about using it [sic] the way most people do; they are talking about leveraging it as a spamming device” (Boyd 2008). Further, Hestres has examined powerful advocacy methods employed by activist groups online (Hestres 2014). Compounding these deliberate attempts to “drown out” opposing voices online are the sampling biases of different social media systems. For example, the, currently popular, Twitter site is estimated to be used by only 16% of Americans (Duggan & Brennan 2013), and reactions measured on Twitter often differ from those measured by surveys (Mitchell & Hitlin 2013). Therefore, if social media are to be deployed as a tool in public consultation about bioenergy developments, serious
consideration is required of how communications can be monitored to gauge public opinion accurately. An industry has arisen to provide tools for social media monitoring, or ‘listening’ (Smith et al. 2014). Some authors claim there are as many as 200 such tools currently available (Stavrakantonakis et al. 2012). These include a mixture of free tools, which are typically single purpose, and multifunctional commercial software (Laine & Frühwirth 2010). Sentiment analysis is a core feature of such tools, important as it is in the analysis of perceptions of organisations, projects and or its products.

Experiences from two analogous fields are pertinent to the particular case of public consultation. The first is scientific communication. In the interests of promoting public engagement with science, researchers have, been encouraged to use social media to communicate about their research (Groffmann et al. 2010). Measuring the attention their social media communications receive (Tortelainen & Katvala 2012) has become one component in the estimation of impact – a critical measure of research value in the prevailing funding climate. The second field of interest is marketing. Marketers have developed a multitude of metrics to assess the effectiveness of social media campaigns, of which those metrics which Barger & Labrecque (Barger and Labrecque, In Press) associate with long-term marketing communication objectives (improving customer satisfaction, creating awareness, building relationships and fostering community) are relevant to the public consultation agenda. Key among these are attitudinal measures, used in marketing to quantify the impact of advertising. These relate to social media metrics such as volume, engagement and number of advocates. Underlying the computation of these metrics is sentiment analysis, since it is necessary to distinguish positive and negative mentions. It is for this reason, because it underlies so much of the measurement of perception and reputation, that we focus on sentiment in this short paper.
In marketing terms sentiment might be used to monitor customer satisfaction with a brand or product. In the public consultation field the analogous use is to monitor public opinion concerning proposed developments. For example, Gao et al. (2014) estimate Twitter users’ attitudes to controversial topics, including a dataset on fracking, by looking at sentiment along with opinion and likelihood to take action, such as spreading a link to a petition. Gao’s work identified eight distinct types of opinion about fracking, which two were classified as positive (Economy & Energy, Safety) and six as negative (Oil Spill, Environment, Health, Economy, General, and Call for Action). This kind of insight into stakeholder attitudes is valuable to companies wishing to ensure that applications address reasonable concerns. The key technology for gaining such insights is sentiment analysis. In the next section, we provide an introduction to research on sentiment analysis for social media, which outlines some of the particular challenges that need to be taken account of to obtain accurate sentiment data from these kinds of texts.

**Sentiment Analysis for Social Media**

Sentiment analysis, also known as opinion mining, aims to determine the attitude of a writer with respect to some topic in text. A basic task in sentiment analysis is identifying the polarity (either positive or negative) of a given text. This can be extended to classify a text into one of the emotion categories, such as Anger, Disgust, Fear, Joy, etc. Other sentiment analysis tasks include retrieving opinions of relevance to a specific topic or query, summarising opinions over multiple text sources towards a certain topic, identifying fake or untruthful opinions, tracking sentiment and topic changes over time, predicting people’s behaviours, market trends, political election outcomes, etc., based on opinions or sentiments expressed in online content.

Sentiment analysis on social media poses new challenges compared to that on conventional text, mainly due to short text length, irregular and ill-formed words, and constant language...
evolution. Previous work on Twitter sentiment analysis relies on machine learning approaches trained on noisy labels. For example, emoticons such as “:), :D, :-(“ and hashtags such as “#fun, #happy, #scary” are taken as the indication of tweet sentiment to train classifiers, which learn a general rule that maps input tweets to a sentiment class (Pak & Paroubek 2010; Purver & Battersby 2012; Suttles & Ide 2013). Obviously, the assumption that emoticons or hashtags are accurate sentiment indicators of tweets is problematic.

Other work explores the use of pre-built lexicons of words weighted with their sentiment orientations to determine the overall sentiment of a given text. For example, Bollen et al. (2011) detected the emotional states in tweets such as “Calm”, “Alert”, “Sure” etc. based on the Profile Of Mood States (POMS) lexicon (Norcross et al. 1984) for stock market prediction. Thelwall et al. (2012) built a human-coded lexicon of words and phrases for social data and used it for the identification of both the polarity orientation and the strength of polarity on the social web. Saif et al. (2014a, 2014b) argued that words’ sentiment orientation and/or sentiment strengths could change depending on context and targeted entities. They proposed an approach which updates words’ sentiment orientations and strengths based on other words co-occurred with them. Their approach outperforms that in (Thelwall et al. 2012) for tweet-level sentiment classification.

In recent years, there has been increasing interest in employing social relations for both document-level and user-level sentiment analysis. It is based on a hypothesis that users connected with each other are likely to express similar opinions. In Twitter, for example, social relations can be established by the following links, through retweeting, or by referring to other users in one's messages using “@” mentions. Speriosu et al (2011) constructed a heterogeneous network that has users, tweets, words, hashtags, and emotions as its nodes, which are connected based on the link existence among them. Sentiment labels were propagated from a small set of nodes seeded with some initial label information throughout
the network. Instead of tweet-level sentiment classification, Tan et al. (2011) incorporated both textual and social relations revealed by the following links and "@" mentions for user-level sentiment detection. Starting from some seed user nodes labeled as positive or negative, they proposed a learning method to propagate sentiment label to all the users in the heterogeneous network consisting of both users and tweets as nodes. Hu et al. (2013) observed sentiment consistency (the sentiments of messages posted by the same user are likely to be consistent) and emotion contagion (sentiments of messages posted by friends are likely to be similar) on Twitter. They proposed to incorporate sentiment relations between tweets into a machine learning approach for tweet-level sentiment classification. While employing social relations for sentiment analysis on social web has shown promising results, building heterogeneous networks capturing such social relations is not an easy task due to the dynamic nature of social networks and incomplete information one can access.

A large body of research work in sentiment analysis from social media still focuses on sentiment detection at the document-level. However, it is not uncommon for users to express mixed sentiments in their online messages. As such, it is crucial to be able to identify sentiment at the, more fine-grained, topic-level. Previous studies (Lin et al. 2012, He et al. 2013) have shown success in simultaneous detection of both topics and topic-associated sentiments from product reviews. However, the performance of such approaches deteriorates when porting to Twitter largely due to the short length of tweet messages. As such, new tools or approaches capable of detecting topic-level sentiment in social media posts are required. In the next section, we illustrate the capacity of sentiment analysis by providing preliminary sentiment analysis of a sample of tweets.

**Entity Level Contextual Sentiment Extraction**

The tweets used in this analysis were collected using the Twitter API. The Twitter stream was
filtered using the term ‘bioenergy’ to identify tweets containing ‘Bioenergy’, ‘bioenergy’, ‘#Bioenergy’ etc. A total of 732 tweets were harvested between the 17th of March and the 2nd of April 2015. The first 300 tweets were examined by two bioenergy experts, who identified entities of interest in the text. These entities provide a proxy for stakeholders, who would be identified in analysis for public consultation. A total of 55 named entities were identified in this way. These included organizations such as @ebri_uk, the European Bioenergy Research Institute, companies and plants, such as hadfields and drax, hashtags, such as #co2-to-fuel, and some individuals.

Data Filtering was applied to the dataset to reduce the amount of noise in the tweets by applying a series of pre-processing steps, including removing duplicate tweets, retweets and non-ascii characters, revert words that contain repeated letters to their original English form (e.g., "loooovve" will be converted to "love"), process contraction and possessive forms (e.g., "he’s" -> "he is"). Applying data filtering resulted in reducing the number of tweets in the dataset to 441 tweets, and the number of entities to 43.

We extract the contextual sentiment of the 43 named-entities detected in our Twitter dataset. To this end, we use the SentiCircle approach (Saif et al. 2014a), which detects the contextual sentiment of a word or an entity from its co-occurrence patterns with other words in tweets. In particular, SentiCircle represents each entity e in a tweet collection T as a vector \(c = (c_1, c_2, ..., c_n)\) of terms that occur with e in any tweet in T. The contextual sentiment of e is then extracted in two steps. First, the context vector c is transformed into a 2d circle representation, where e is positioned at the center of the circle, and each point around it represents a context term \(c_i \in c\). The position of \(c_i\), as illustrated in Figure 1, is defined jointly by (i) an angle (\(\theta\)) representing the prior sentiment orientation of \(c_i\) and (ii) a radius (\(r\)) representing the degree of correlation between \(c_i\) and the entity e.

The trigonometric properties of the SentiCircle allow dividing the circle into four sentiment
quadrants as shown in Figure 1. Terms in the two upper quadrants have a positive sentiment ($\sin \theta > 0$), with the upper left quadrant representing stronger positive sentiment since it has larger angle values than those in the top right quadrant. Similarly, terms in the two lower quadrants have negative sentiment values ($\sin \theta < 0$). Moreover, a small region called the “Neutral Region” can be defined. This region is located very close to X-axis in the “Positive” and the “Negative” quadrants only, where terms lie in this region have very weak sentiment (i.e., $|\theta| \approx 0$).

![SentiCircle of an entity e](image)

**Figure 1 SentiCircle of an entity e**

The overall contextual sentiment is then calculated by extracting the geometric median of the points (context terms) within the circle. The position of the median within the circle represents the overall contextual sentiment of $e$. i.e., the sentiment of $e$ is considered positive if the median lies in positive quadrants, negative if the median lies in the negative quadrants, and neutral if the median lies in the neutral region.

**Evaluation Results**
Figure 2 shows the contextual sentiment (positive, negative, neutral) of the 43 named-entities in our datasets calculated using the SentiCircle approach, as explained in the previous section. According to these results, 18 entities occur with a neutral sentiment in the dataset, 14 entities occur with positive sentiment and 11 entities occur with negative sentiment.

Figure 2 Contextual sentiment distribution of the 43 named entities in the Twitter dataset

Figure 3 shows the SentiCircle of the Twitter bioenergy dataset. Points inside the circle denote the 43 named entities and are positioned based on the median of their own SentiCircles. As we can see, entities, such as @businessgreen and hadfields, receive a positive sentiment since their median points are positioned in the first and second quadrants in the SentiCircle. These entities typically represent research organizations, such as EBRI and ETI, as well as trade magazines such as Business Green and Utility week. We would expect these to take a positive view of bioenergy. Entities, such as waste, and @bioenergyintl are positioned in the negative quadrants, receiving therefore a negative sentiment. These quadrants contains known opposition groups such as @biofuelwatch.
Entities like @aebiom, and @mitglobalchange have a neutral sentiment as their medians lie in the neutral region close to the origin of the SentiCircle. Interestingly this group also includes the opposition group @climatejustice.

![SentiCircle of all the 43 entities in Twitter dataset.](image)

**Figure 3 SentiCircle of all the 43 entities in Twitter dataset.**

Lastly, Table 1 lists all the 43 entities under each sentiment class.

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>@ebri_uk</td>
<td>@centerforbiodiv</td>
<td>@doe_jgi</td>
</tr>
<tr>
<td>@the_eti</td>
<td>#nbbio2015</td>
<td>@mitglobalchange</td>
</tr>
<tr>
<td>bioenergy insight magazine</td>
<td>@bioenergyintl</td>
<td>aebiom</td>
</tr>
<tr>
<td>@utilityweek</td>
<td>axioma</td>
<td>avebiom</td>
</tr>
<tr>
<td>@businessgreen</td>
<td>saxlund</td>
<td>@aebiom</td>
</tr>
<tr>
<td>hadfields</td>
<td>@alddyendonnelly</td>
<td>@foresteurope</td>
</tr>
<tr>
<td>@suzannewaldman</td>
<td>@neil1808</td>
<td>drax</td>
</tr>
<tr>
<td>@johndpmorgan</td>
<td>@sashalyutse</td>
<td>billington</td>
</tr>
<tr>
<td>@arthurhcyip</td>
<td>@biofuelwatch</td>
<td>@purplenergy</td>
</tr>
<tr>
<td>tilbury project</td>
<td>@ran</td>
<td>@stollmeyereu</td>
</tr>
<tr>
<td>#ecosystemservices</td>
<td>waste</td>
<td>@arielbrunner</td>
</tr>
<tr>
<td>#beyondbiomass</td>
<td></td>
<td>@climatejustice1</td>
</tr>
<tr>
<td>#co2-to-fuel</td>
<td></td>
<td>land-use legacies</td>
</tr>
<tr>
<td>@idiottracker</td>
<td></td>
<td>drax</td>
</tr>
<tr>
<td></td>
<td></td>
<td>#pyrofab</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bioenergy crops</td>
</tr>
<tr>
<td></td>
<td></td>
<td>#bioenergy development</td>
</tr>
</tbody>
</table>
Table 1. 43 entities under each sentiment class

Conclusion

There are clear real world implications of the negative sentiment that community stakeholders and opposition organisations often express towards energy projects. There have been several studies that have reviewed the key stakeholder barriers to project development (Adams, Hammond et al. 2011; Wright, Dey et al. 2014), and more specifically the effects of local opposition (Rösch and Kaltschmitt 1999; Upreti and van der Horst 2004). Rösch and Kaltschmitt (1999) categorise the key concerns as: traffic, local employment, local and regional environment, attractiveness and image of the community. It is increasingly common for developers of projects to financially compensate the local community of a prospective development in each of these areas. The Economist (2013) reports that although there isn’t a standard way or minimum sum for compensating the local community in the case of onshore wind turbines developments, the money typically goes towards schools, village halls or to reduce the utility bills of the local residents. The Scottish Government (2014) published a ‘good practice’ guide for the remuneration of the local community, which gives a guidance figure of £5k per MWel capacity per annum for the life of the project.

These real world costs of negative sentiment for bioenergy and EfW project development thus make the consultation process and the measurement of sentiment imperatives for businesses operating in the bioenergy market. Public participation research demonstrates that social media should be utilized to a greater extent to improve information flows in the consultation process. If this were implemented data could be obtained providing a valuable
resource for businesses and policy makers. The SentiCircle analysis presented here demonstrates that sentiment analysis provides promising insights: with entities which post positively about bioenergy identified by experts as including researchers and trade magazines, and with entities which post negatively identified as including opposition groups. When scaled up to larger samples, collected with respect to specific developments, this technology has the capacity to provide better understanding of public opinion as part of the consultation process.

Acknowledgments

Daniel Wright & James Scott are funded by the BioenNW (INTERREG IVB) project, This work was supported by the EU-FP7 project SENSE4US (grant no. 611242).

References


Hestres, L. (2014). “Preaching to the choir: Internet-mediated advocacy, issue public mobilization, and climate change”. *New Media & Society*, 16(2) 323-339.


About the authors

Victoria Uren lectures in Information Systems at Aston Business School and has previously held research positions at the University of Sheffield and the Open University. Her research interests are in semantic knowledge management, how people engage with scientific discussion on the web in social media, and the use of games and gaming in business and education. She received her PhD on text categorisation in 2001 from Portsmouth University (UK).

Daniel Wright is a Research Associate within European Bioenergy Research Institute (EBRI) specialising in the application of operational research techniques to support decision-making within the bioenergy sector. He completed his PhD (2013) in the lifecycle planning and risk management of small-scale biomass combined heat and power (CHP) project development and currently works on the BioenNW (INTERREG IVB) project.

Jim Scott works as a Research Associate for the European Bioenergy Research Institute (EBRI) on developing small to medium scale sustainable energy from waste projects. He completed his PhD (2013) on strategic sourcing and purchasing of biomass for UK combustion power projects and currently works on the BioenNW (INTERREG IVB) project.

Yulan He joined the School of Engineering and Applied Science, Aston University, UK, as a Senior Lecturer in 2013, having previously held academic positions at the Open University, University of Exeter, University of Reading, and Nanyang Technological University, Singapore. Her main interests are in social media analysis; sentiment analysis and opinion
mining; business intelligence and data mining; text mining; machine learning; and bioinformatics.

Hassan Saif is currently a research associate at the Knowledge Media Institute (KMi), The Open University. He obtained his PhD in sentiment analysis and opinion mining from the Open University in 2015. Before joining KMi, Hassan worked as a research assistant at the University of Lincoln, UK and as a research engineer for Nordic Sense, Finland. Hassan’s research interests are focused on Semantic Sentiment Analysis of microblogs, and more specifically on how semantics can be used to enhance current Sentiment Analysis models. Hassan has published a series of articles about this topic in various leading conferences and journals (WWW, ISWC, ESWC, LREC) and his research has driven the development of novel sentiment analysis technology for two EU funded projects (ROBUST and Sense4us).