

Sustainable Agriculture through ICT innovation

The representation of Food in the Social Media Forum TwitterElizabeth York¹ and Christopher Brewster¹¹Aston Business School, Aston University, Birmingham, UK B7 4ET
{E.York|C.A.Brewster}@aston.ac.uk**ABSTRACT**

Social Media is becoming an increasingly important part of people's lives and is being used increasingly in the food and agriculture sector. This paper considers the extent to which each section of the food supply chain is represented in Twitter and use the hashtag *#food*. We looked at the 20 most popular words for each part of the supply chain by categorising 5000 randomly selected tweets to different sections of the food chain and then analysing each category. We sorted the users by those who tweeted most frequently and categorised their position in the food supply chain. Finally to consider the indegree of influence, we took the top 100 tweeters from the previous list and consider what following these users have. From this we found that consumers are the most represented area of the food chain, and logistics is the least represented. Consumers had 51.50% of the users and 87.42% of the top words tweeted from that part of the food chain. We found little evidence of logistics representation for either tweets or users (0.84% and 0.35% respectively). The top users were found to follow a high percentage of their own followers with most having over 70% the same. This research will bring greater understanding of how people perceive the food sector and how Twitter can be used within this sector.

Keywords: Social media, supply chain, consumers, logistics, UK**1. INTRODUCTION**

It may be obvious that food is an important part of people's lives; however, its presence in the technology we use is not so obvious. Food works for us on so many levels – it is fuel for all of us, it can be a hobby, it can indicate where we have been in the world, what we have experienced, it can also be a point of pride (Costello, 2011). Food is also an occasion for sharing, for distributing and giving, for the expression of altruism, whether from parents to children, children to in-laws, or anyone to visitors and strangers (Fox, 2012). Food plays a major role as an expression of people's culture and identity (Fischler, 1988; Scholliers, 2001), something apparent in the close association of food with particular ethnic groups, countries or regions of the world. Thus we have Italian restaurants, Thai food, and Moroccan cuisine. Consequently, it is hardly surprising that food, as a topic, has a substantial presence on the World Wide Web in a variety of manifestations. It is not surprising that food and agriculture have found a home in social media. This paper is an initial contribution to assessing the presence of different food sectors on the social networking site Twitter¹.

¹ <https://twitter.com>

Sustainable Agriculture through ICT innovation

Over the past decade, the Internet has changed the way people work, play, learn, and communicate (Leung and Lee, 2005; Shirky, 2008) and this has had a large impact on our relationship with food. Food is very much a part of popular culture, and the beliefs, practices, and trends in a culture affect its eating practices. Popular culture includes the ideas and objects generated by a society, including commercial, political, media, and other systems, as well as the impact of these ideas and objects on society (Rodriguez, 2012). Consequently, instead of turning up to a restaurant unprepared, we can now check their reviews beforehand on-line. Instead of cooking what we imagine is traditional cuisine; we can ask a resident of the area on a recipe site. Instead of wondering what the nutritional contents of a chocolate bar may be, one can tweet the company. Instead of children believing milk comes from chickens, we can involve them in running a farm (Manchester Evening News, 2007). Social networking has opened a whole new way for people to interact with food. Blanchfield (1996) found that there are substantial opportunities both to acquire and to communicate food information, and that it may be expected to proliferate and expand as fast as the Internet itself.

Twitter is a popular micro-blogging social network that was founded March 21, 2006 (Pochampally and Varma, 2011; Krishnamurthy et al., 2008). Twitter restricts the user by only allowing them to use 140 characters in the message called a “tweet” (Westerman et al., 2012). The tweets can be posted in many different ways; via the web, text messages, or instant message, and most frequently now via a client on a smartphone (Pochampally and Varma, 2011). Topics range from daily life to current events, news stories, and other interests. Twitter users will both follow other users or themselves be followed; a user with a greater number of followers is seen as being knowledgeable and creditable (Westerman et al., 2012; Java et al., 2007).

Due to our love of food, it did not take long for food related organisations to join social media. There are representatives from all areas of food, big and small retailers, policy creators, educationists, food establishments, producers, food lovers and to a lesser extent the food industry. No one to our knowledge has previously considered how social media is being used in the area of food.

In this paper we will be considering the behaviour of Twitter users in relation to food and how food is discussed on the social media site Twitter. Through this we will be looking at: 1) What are the most common words relating to food and how do they cover the food chain? 2) Who are the most common tweeters relating to food and do they come from specific areas of the food chain? 3) What percentage of these tweeters are individuals? 4) What percentage of these tweeters are individuals? 5) What is the influence of these users? The rest of this paper is structured as follows: first, we present our methodology and explain how the data we used was collected. Then we review the results, followed by a discussion of the results, and end with a brief conclusion.

Method

Data Collection

The data was collected using the Twitter API and a Python script, querying only for

Sustainable Agriculture through ICT innovation

Tweets containing the *#food* hashtag. Data was collected over a 4 day period in January 2013. Once the collection period had ended, the data was analysed using a combination of scripts written in the statistical programme R and in Python. The users and tweets were categorised separately by two human coders. The text was split into five main areas as *agriculture*, *logistics*, *retail*, *consumer* and *society*. There was also a 6th category for tweets that were undecipherable due to language barriers. *Agriculture* was defined as those tweets that discuss the growing or production of food. *Logistics* was defined as the transporting, packaging, storage and processing of food items ready for sale. *Retail* was the sale of items or the establishments at which items are sold, this includes eating establishments. *Consumers* are defined as those tweets that are post-retail, referring to the consumption or cooking of the food. *Society* is defined as tweets that discuss news articles and global issues. The users, in addition, were split into five categories, *agriculture*, *logistics*, *retail*, *foodies* and *consumers*. A 6th category again was used for accounts that were protected, suspended or undecipherable. *Agriculture*, *logistics* and *retail* had the same meaning as above. *Foodie* is an informal term for a particular group devoted to food and drink who would tweet multiple times on this topic. *Consumer* in this instance is slightly different as this here it refers to users that tweet about food but do not do so continuously or refer to food in their personal description. The users were split into a second set of categories depending on whether the account was used by an individual or contributed to by multiply people i.e. they were an organisation. For example a business or volunteer group would be classed as an organisation. Intercoder reliability was checked (Kappa) and the percentage agreements were 81.43% for the text and for the user categories, and for user types the result was 81.40% and 82.56% respectively.

The frequency of terms was calculated for each tweet category being analysed to identify the top 20 most common words. Firstly for each stop words were removed before a word count was undertaken. The top 10 users were also analysed for their influence by collecting data on the number of followers and people they followed and comparing the overlap two lists. The connection between the top 5, 10, 25, 50 and 100 users was analysed by looking at the number of occurrences of a user within either the people followed or the followers group.

Results

Tweets

Of the 91312 tweets collected we randomised the results and analysed 5000 tweets. We therefore considered 5.48% of the collection though 444 were considered undecipherable and removed from the analysis. The tweets were split into five categories according to their subject matter (Table 1) as described above. *Consumers* had the highest number of tweets with 3983 of the 5000 analysed. Tweets assigned to the category *consumers* if they discussed food but had not direct association to the other categories, for example, "Trying some almond milk for the first time. #food <http://t.co/rc5Rrsou>" and "Ate too much food!! @devkip04 #icco's #food #bedtime". *Logistics* was assigned the lowest number of tweets with only 16 of the 5000 analysed. The tweets categorised as *logistics*



Sustainable Agriculture through ICT innovation

mainly discussed transport, manufacturing and packaging, for example “KM Packaging is chosen by global food leader - #Technology #Flexibles #Food <http://t.co/cZdYtveR>”. *Society* was an added category as when analysing the tweets as a number discussed global issues, for example food hunger, waste and sustainability. This category was assigned to 96 tweets, including the tweets “Thirteen Resolutions to Change the #Food System in 2013: <http://t.co/4s0VxtNK> #health” and “One billion tons of #food produced by #human consumption goes to waste around the #world each year.” 74.89% of the tweets included links to photos and further information for instance news articles and blogs. This suggests that although tweets can only be 140 characters, people like to give further explanations and information.

Table 1: The number of tweets assigned to each category and the percentage of the number of tweets analysed per category

Code	All	Percentage
Agriculture	39	0.86
Logistics	16	0.35
Retail	422	9.26
Consumer	3983	87.42
Society	96	2.2
Unusable	444	
Total	5000	100

Term Frequency

The top twenty ranked terms from each category (i.e. some times more than 20 words) were selected, and these were combined giving a total of 92 different words. Only three words appeared in 3 or more categories viz. *safety*, *year* and *healthy*. *Healthy* appeared in 4 categories, as it did not appear in *Logistics*. All top three words in frequency appeared in the *Agriculture* and *Society* categories.

Though logistics had the fewest number of tweets, when normalized, its terms show greatest consistent usage (Figure 1). Whether this would occur if more tweets were analysed cannot be said. There were 14 words that appeared in 2 or more categories suggesting there is not a common set of words used for describing food and the supply chain. However, there is an overlap between the categories with *agriculture* and *society* having a large number of the same words and *retail* and *consumers* as well. This would be expected of the retail and consumer as these categories are very similar, it is less expected of agriculture and society but as much of the public discussion is about hunger, sustainability etc. then it is less surprising. This suggest that there is not a common set of words used to describe food but different sets depending on the part of the supply chain and that overlap is minimal between the categories.

Sustainable Agriculture through ICT innovation

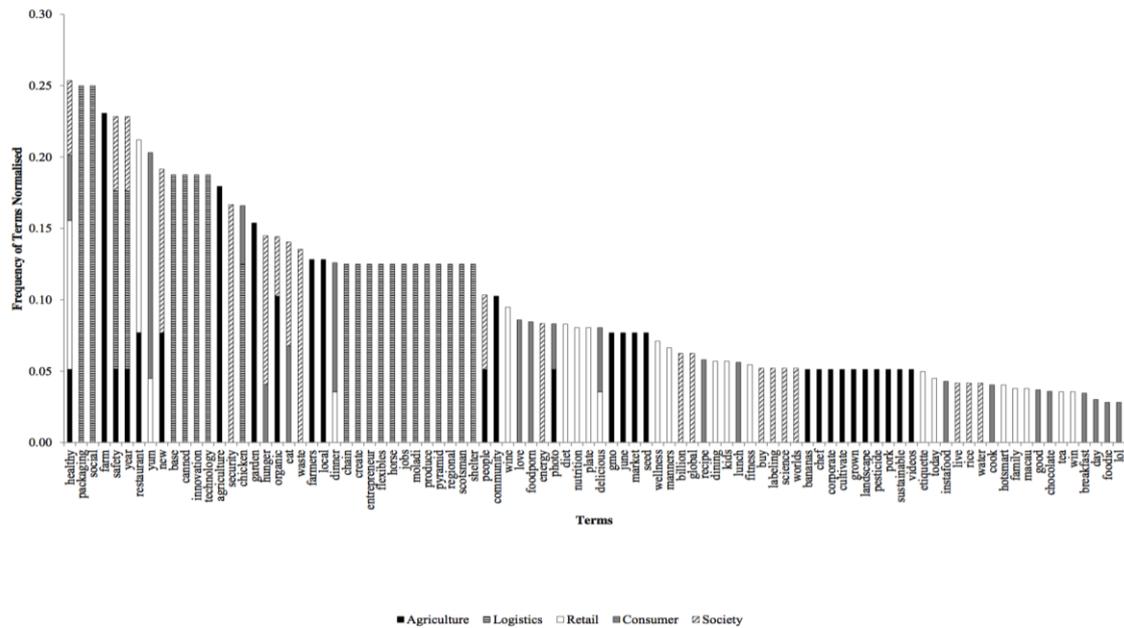


Figure 1: The top most common frequency terms for each subject area combined and normalised

Users

In the total of 91312 Tweets collected, a total of 65303 different users were identified. We focused on the 852 (1.30%) users who tweeted 6 times or more. This resulted in 15299 tweets, 16.75% of the tweets collected. The users were split into five categories (as described above) with a sixth being used for unusable accounts (i.e. in a foreign language or the accounts had been closed). The five categories were *agriculture*, *logistics*, *retail*, *consumer* and *Foodies*. As with the tweets, the *logistics* category had the least number of users, 7, and the *consumers* the greatest number, 428. *Agriculture* was the second smallest category with 13 users and *retail* and *foodies* had 172 and 211, respectively. Of the top 10 users for all categories, 4 came from *retail*, 3 were *consumers*, and 3 were classified as *Foodies*. Retail having 4 of the top 10 user is not unexpected as retail would tend to tweet more to keep users interested in their products and services. An analysis of the type of user was also conducted. Where a user represents themselves as an organization, the type of tweets they sent and the amount they tweeted was different. 6 of the top 10 users were found to be organisations. For the rest of the users analysed, it was found that 73.40% were individuals and 26.60 % were organisations, having 610 and 221 users respectively.

The ids of the followers and people they followed of the top 10 users were compared to find the percentage of the same users in both groups. Interestingly, the users ranked 6-10 have a great overlap of followers and people they follow of over 88%. In contrast, the top five users are mixed in whether they have more followers or people they follow. For the first user, *drdaveanddee*, they have many more people following them than that they are following (91.12%). This is similar with *searedfoodblog* and *breakfastmag*.

Sustainable Agriculture through ICT innovation

Users *michael61211* and *askauntbettycom* have more people they follow than people who follow them. These differences have implications for the import and impact of the tweets they produce.

To consider the number of linked users discussing food we looked at the top 5, 10, 25, 50 and 100 users to see how many of the users were the same in the group for both friends (ds) and followers (ers). The followers and friends followed a similar pattern for all 5 groups and a similarity between friends and followers for each group. There is a general pattern that if you double the number of users being analysed the number of users that are the same increases by 5.

Discussion

We firstly split the tweets and looked at the number of tweets in each group for the categories that span the supply chain. The most common category was consumer which is unsurprising when you consider that everyone must eat and as it is an important part of life people are likely to discuss in some form with friends. The least common category was logistics probably due there being a small number of actors in that sector. More companies are concerned with the movement of foodstuff rather than processing, and the use of twitter in this area is very low. There is great possibility for its use in logistics as discussed by (Logistics Business, 2011) however these possibilities could only occur with increased trust and cooperation in the sector.

Agriculture is a growing area of twitter with ministry departments having accounts and guides for member of the community on how to use twitter (Cho and Park, 2012; Futures, 2010). Therefore it was surprising there were not more with only 0.86% of tweets analysed being agricultural based, however the top 3 words were agriculture, garden and farmers. Here we need to signal an obvious limitation of the study in that we have only focused on tweets containing the hashtag *#food* and we are aware of a great deal of activity in the agriculture domain under different hashtags e.g. *#agchat*, *#agchatuk*, etc.

Society was a category we added to the analysis as we found that many tweets were appearing that discussed sustainability, science, hunger and waste and felt that these were being tweeted to raise awareness to the community on the bigger issues that society faces today. Considering the most frequent words for this category were security, waste and news we can start to see the issues that are concerning users today.

From our user analysis as with the tweet analysis we found that the number of member of the logistics part of the supply chain is very low. This again could be due to the reason given for the tweets, however, another reason would be due to there being no need for the logistics users to tweet. Many of these users would probably find customers through other avenues and therefore do not feel the need to advertise their presence to the community; it may also be that people are simply just not interested in the processing, packaging and transporting of their food.

Foodies were a category added to the analysis due to the need for some way to differentiate general people tweeting about food and users who have a passion for food. Most

Sustainable Agriculture through ICT innovation

foodies tended to tweet very frequently using the #food and mentioned it as part of their profile description. These people differ from a person who happens to mention food in a long series of tweets that described their day as a whole. Foodies tended to cover many food topics in their tweets ranging from news articles to recipes.

Surprisingly not many users were retail based. With the ever increasing use of social media in marketing then the number of retailers present was lower than expected, however 4 of the top 10 users were retailers suggesting that a few use Twitter frequently rather than many using it sometimes. For the retailers to have such a small presence could be to do with them concentrating on their own hashtags more.

The top ten users did not have an exceptionally large set of people they follow and followers, with most having less than 5000. They did have a large overlap of people they follow being follower and vice versa suggesting that they tend to follow those following them. This was apparent for all except the top four who tended to have more followers suggesting that they had a greater influence on the community. Their tweets were more likely to travel further than their local network. However the first and third top user had less followers cutting down on their impact to influence others as though they might tweet often they do not spread their message to many. We did consider how connected the user were by comparing the group of friends and separately the groups of follower, there was some connectivity which increased as the group increased in size, but was not significantly large, this is probably due to the number of people involved. If you studied a more specialist hashtag then the likelihood on connectivity would increase.

Conclusion

Food is a significant part of life and has a large presence on twitter. Most information about food is communicated through article links which lead though to other sites. Twitter is mainly used to pass information rather than to build a community sense. Many people use twitter to discuss food though not many use it to discuss the various sectors of the supply chain with some sector representation being nearly non-existent. It was also found that though there is some overlap in the frequency of the words for different parts of the supply chain but overall it was not large and therefore each sector would have its own vocabulary. With sectors having their own vocabulary there is likelihood that they would use sector specific hash tags rather than a generalised terms such as food. In future it would be interesting to look at the retweets, the type of links and the geographical location of users and tweets.

References

- J Blanchfield. Food science and technology on the internet. Trends in Food Science & Technology, 7:18, 1996.
- Seong Eun Cho and Han Woo Park. Government organizations innovative use of the internet: The case of the twitter activity of south koreas ministry for food, agriculture, forestry and fisheries. Scientometrics, 90:923, 2012.

Sustainable Agriculture through ICT innovation

- Conor Costello. 10 of the world's most popular foods which have you eaten?, June 2011. URL <http://blogs.oxfam.org/en/blog/11-06-16-10-worlds-most-popular-foods-which-have-you-eaten>.
- Farming Futures. How to use twitter, 2010. URL <http://www.farmingfutures.org.uk/resources/how-use-twitter>
- Claude Fischler. Food, self and identity social science information. *Anthropology of Food*, 27:275–292, 1988. Robin Fox. Food and eating: An anthropological perspective, 2012. URL http://www.sirc.org/publik/food_and_eating_1.html
- Akshay Java, Xiaodan Song, Tim Finin, and Belle Tseng. Why we twitter: Understanding microblogging usage and communities. In Joint 9th WEBKDD and 1st SNA-KDD Workshop, 2007.
- Balachander Krishnamurthy, Phillipa Gill, and Martin Arlitt. A few chirps about twitter. In Proceedings of the first workshop on Online social networks, 2008.
- Louis Leung and Paul Lee. Multiple determinants of life quality: the roles of internet activities, use of new media, social support, and leisure activities. *Telematics and Informatics*, 22:161180, 2005.
- Ingrid Lunden. Analyst: Twitter passed 500m users in june 2012, 140m of them in us; jakarta biggest tweeting city, July 2012. URL <http://techcrunch.com/2012/07/30/analyst-twitter-passed-500m-users-in-june-2012-140m-of-them-in-us-jakarta-biggest-tweeting-city>
- Manchester Evening News. City kids 'think cows lay eggs', 2007. URL http://menmedia.co.uk/manchesterevening-news/news/education/s/1000/1000505_city_kids_think_cows_lay_eggs.html
- Ravali Pochampally and Vasudeva Varma. User context as a source of topic retrieval in twitter. In SIGIR 2011 Workshop on Enriching Information Retrieval, 2011.
- Judith Rodriguez. Popular culture, food and current trends, 2012. URL <http://www.faqs.org/nutrition/Ome-Pop/Popular-Culture-Food-and.html>. Accessed 09 January 2012.
- Peter Scholliers. Food, drink and identity: cooking, eating and drink-ing in Europe since the Middle Ages. Berg Publishers, 2001.
- Clay Shirky. Here Comes Everybody: The Power of Organizing Without Organizations. Penguin Press, 2008.
- The Logistics Business Ltd. Tweeting logistics- will social media play a role in the future of supply chain management?, 2011. URL <http://www.logistics.co.uk/component/k2/item/1216-tweeting-logistics-will-social-media-play-a-role-in-the-future-of-supply-chain-management>. Accessed the 20 February 2013.
- David Westerman, Patrick Spence, and Brandon Van Der Heide. A social network as information: The effect of system generated reports of connectedness on credibility on twitter. *Computers in Human Behavior*, 28:199206, 2012.