

Understanding Actionable Knowledge in Social Media: BBC Question Time and Twitter, a Case Study.

Maria Angela Ferrario[†], Will Simm[†], Jon Whittle[†], Paul Rayson[†], Maria Terzi[†], Jane Binner[§]

[†]School of Computing and Communications, Lancaster University, LA1 4WA, Lancaster, UK

[§]Management School, The University of Sheffield, S1 4DT, Sheffield, UK

Abstract

Automatically extracting actionable knowledge from on line social media has attracted a growing interest from both academia and the industry. However, little clarity exists in relation to what actionable knowledge is, whether it can be measured and where it is more likely to be found. This paper makes an attempt at answering the above questions by gaining a better understanding of actionable knowledge in Twitter.

Introduction

Actionable knowledge discovery from user-generated content is a commodity much sought after by industry and market research. From topic clustering (Xiaofeng 2001) to sentiment analysis (Pang and Lee 2008), academic research has produced a range of data mining approaches that are being used by social media analytic services to produce aggregated summaries of on-line conversations. These aggregates are then used by marketing experts to ‘extract’ actionable knowledge, or to put it simply, they are ‘interpreted’ in order to identify potential business strategies. The need of experts shows that these summaries still contain much ‘noise’ or, in other words, they contain comments that cannot be directly acted upon. Our argument is that if we could automatically filter out such ‘noise’ and only present actionable comments, then the decision making process would be easier. We argue that to achieve such a goal we must gain a better understanding of what actionable knowledge is, where it can be found and what kind of language structures it contains. The aim of this paper is to do so by analysing actionable knowledge in on-line social media conversation and, specifically, in Twitter. As a first step, we clarify the notion of actionability: we consider as actionable those expressions (or tweets) that contain a request or a suggestion that can be acted upon. Actionability is henceforth defined as ‘the measure of actionable knowledge in a statement’ (Simm et al. 2010). This notion of actionability builds on the findings from previous interdisciplinary research looking into ways for capturing actionable knowledge through

digital platforms ‘in the wild’ (Whittle et al. 2010). For example, let’s examine the following two tweets:

A. ‘Yes, give power to communities...BUT DON'T TAKE AWAY ALL THEIR RESOURCES FIRST’

B. ‘We’re building a nasty society not big. Horrible’

Both tweets contain an opinion about the same topic (e.g. society/social innovation). However, tweet A suggests two clear actions (‘give power to communities’ and ‘don’t take away their resources’) while tweet B simply shares a negative view on the subject. We argue that tweet A says something ‘more’ than tweet B in that A contains actionable knowledge and B does not. We believe that it is important to understand not only *how people feel* about a topic but also *what actions* they would like to take and to be taken. To this aim, we first identify which types of tweets are more likely to contain actionable knowledge.

Tweet Types and Hypothesis Testing

Three types of tweets are here identified: closed, open and retweeted (Boyd, Golder and Lotan 2010, Honeycutt and Herring 2009; Java et al. 2007). We investigate if actionable tweets can be more frequently found in any of these categories. Closed tweets address one or more Twitter users via the sign @ followed by the username of the intended recipient. Retweets are reposted tweets and they usually contain the string RT followed by @username where ‘username’ indicates the person that posted the tweet in the first place. Open tweets are tweets that do not fall in either of the above categories.

Our first hypothesis tests whether different types of tweets (closed, open and retweeted) contain different levels of actionability. We also argue that there is another type of tweet: the ‘linked tweet’, a tweet that links to a topic via a hashtag (#). This is a conversational style which is unique to Twitter and it is established when users intentionally include the hashtag to both ‘label’ and ‘link into’ a public conversation. The present research focuses on the use of a specific hashtag (#bbcqt) and reports the results of this initial case study. The #bbcqt hashtag is used for BBC Question Time (BBCQT), a current affairs discussion programme broadcast every Thursday on BBC One in the UK. We chose BBCQT for two reasons: (a) it is a long-standing programme with a well established Twitter audience and (b) its choice of topics (e.g. budget cuts,

immigration laws) is designed to fuel discussions amongst the audience. For this, we argue, it should serve as example of an actionable knowledge ‘trigger’. Hence the second hypothesis to test is whether tweets that contain the #bbcqt hashtag (BBCQT tweets) have higher levels of actionability than tweets randomly sampled from general conversation. Finally, we test whether there is a different distribution of tweet types across three datasets (third hypothesis). In summary, we analyse the data in three steps: firstly we manually annotate the three subsets with actionability scores. Secondly we test our hypotheses by performing statistical analysis on the annotated data. We then use WMatrix (Rayson 2008) to automatically identify language patterns in actionable data.

Research Context

Much research exists on the use (Java et al. 2007), usefulness (Agichtein et al. 2008) and quality (Ghosh and McAfee 2011) of user generated-content. Twitter, in particular, has attracted researchers from a very early stage due to its publicly available content and its conversational nature (Zhao and Rosson 2009). Much research also exists on automated content analysis in the field of sentiment analysis (Hopkins and King 2010) and topic extraction (Cataldi, Di Caro and Schifanella 2010). In addition, the concept of actionability is not a new one: areas of research include knowledge discovery (Cao and Zhang 2006), opinion mining (Pang and Lee 2008) and crisis communication (Neubig et al. 2011). However, very few have so far attempted a clear definition of actionability (Simm et al. 2010) or focused on its structures and the context of its exchange. In this study, we focus on the dynamic aspect of the language as we aim to investigate how people ‘use language to bring about change’ (Thomas 1995).

Statistical Analysis

We analyse three datasets of a combined total of 941 tweets extracted from a full dataset of 2 million tweets posted from the British Isles over a period of three weeks (24th January – 13th February 2011). During this period three different BBCQT programmes were broadcast: the first on 27th January, the second on 3rd February and the third on 10th February.

The first dataset – General Conversation – contains 350 tweets randomly selected from the full set. Firstly, each tweet is automatically annotated with the three tweet types (open, closed, retweeted). Each record is then manually annotated with an actionability score using a 3-point Likert scale (1=‘not actionable’, 2=‘somewhat actionable’, 3=‘actionable’) by three researchers (kappa-agreement 0.83). The second dataset – Sampled BBCQT – consists of 304 records. These tweets were randomly sampled (95% CL, 5% CI) from all the 1,452 tweets containing #bbcqt and posted during the three-week target period. Finally the third dataset – 10thFeb BBCQT – contains the full set of

tweets posted on 10th February, an episode randomly chosen from the three BBCQT broadcasts. This dataset contains all the 310 BBCQT tweets posted on 10th February except the ones that were already included in Sampled BBCQT.

Summary of Results

Overall 30% of the tweets have been annotated as either ‘actionable’ (18%) or ‘somewhat actionable’ (12%). This finding indicates that Twitter users share actionable knowledge in a proportion that could be advantageous for extracting a first-cut of data when mining large datasets. This would allow the identification of topics for which the general public feels strongly enough not only to share an opinion but also to recommend actions. We test our first hypothesis to investigate whether different types of tweets contain different levels of actionability. The *Pearson Chi square* test finds no statistically significant relationship ($p > 0.05$) between Tweet Types and Actionability. Table 1 provides a summary of results for the other Hypotheses: 2a, 2b and 3.

Table 1 Summary of Results for Hypotheses 2a, 2b, 3

	Pearson χ^2 (p-values)	Results
2a	.018	<i>General Conversation</i> contains a significantly higher proportion of actionable tweets than <i>Sampled BBCQT</i>
2b	.000	<i>10thFeb BBCQT</i> contains a significantly higher proportion of actionable tweets than <i>Sampled BBCQT</i> and <i>General Conversation</i>
3	.000	<i>Sampled BBCQT</i> and <i>10thFeb BBCQT</i> contain a significantly different distribution of tweets types to <i>General Conversation</i>

Discussion

The statistical analysis holds some initial interesting results. Firstly, it shows that Twitter users spontaneously share actionable knowledge in daily conversation and that they do so in both open and closed tweets (Hypothesis 1). Secondly, the tests show that Sampled BBCQT contains significantly less actionable knowledge than General Conversation (Hypothesis 2a), but also that different episodes (e.g. 10thFeb BBCQT) may attract more actionable comments than others (Hypothesis 2b). We suggest that a number of factors can impact on actionability of the BBCQT tweets including (a) the posting time of the tweets, (b) the topics of discussion and (c) the panellists. A further investigation finds (a) no statistically significant association (at $p < 0.05$) between actionability levels and posting times. As for the other two factors, namely (b) the topics of discussion and (c) the choice of the panellists, findings from the automated content comparison reported below suggest that they are associated with actionability and they will be therein discussed. Finally, Hypothesis 3 shows that there are significantly more open tweets in both BBCQT datasets

than in the random sample suggesting that Twitter users tend to engage with the audience of a programme as a whole rather than with specific recipients.

Actionable vs. non-Actionable Data

In this section we wish to investigate the language differences between actionable and non actionable tweets in order to identify what we refer to as *actionability markers*, language expressions that are more frequently found in actionable data than in not actionable data. We use Wmatrix (Rayson 2008) to compare two sets: Set_a containing all the actionable data extracted from the three datasets (171 tweets); and Set_{na} containing all the non actionable data (671 tweets). Wmatrix is a web-based software tool for corpus analysis and comparison. It is built on the CLAWS and USAS corpus annotation tools, which tag each word in the text with its Part Of Speech (POS) and a semantic category (SemTag). Given the two sets, we then use Wmatrix to produce three types of frequency lists: Word, POS and SemTag.

Method

Two sets are prepared: Set_a containing actionable data (actionability score = 3) and Set_{na} containing not actionable data (actionability score = 1). We exclude data with actionability score of 2 since we are only interested in comparing data of clear or no actionability. The two sets are then loaded in Wmatrix which tags the tweets with POS and SemTags. Three frequency lists are then compared using the log-likelihood (LL) test. This process is carried out in two parts: in the first iteration we use Set_a as normative corpus; Set_{na} is used as normative corpus in the second iteration.

Discussion

Table 2 and Table 3 show the Keyword and the SemTag frequency lists for Set_a calculated in the first iteration. In both tables O_a is the observed frequency in Set_a and O_{na} is the observed frequency in Set_{na}. The relative frequencies in the texts are shown as %_a and %_{na} and sorted on LL value to show key items at the top. We set the statistical significance to the $p < 0.01$ (LL > 6.63). Due to space restrictions we only report the top-ten scores.

Table 2 Key Word List: Iteration I

Item	O _a	% _a	O _{na}	% _{na}	LL
should	28	1.2	5	0.07	52.29
need	12	0.51	0	0	33.2
please	10	0.43	0	0	27.67
take	6	0.26	0	0	16.6
better	7	0.3	1	0.01	13.92
get	17	0.73	13	0.19	13.49
jobs	4	0.17	0	0	11.07
shut	4	0.17	0	0	11.07
to	60	2.56	105	1.5	10.32
would	10	0.43	6	0.09	9.96

Table 2 shows an overuse of modal verbs expressing duty (should), necessity (need), requests (please), commands (let's) and conditions (if). Table 3 contains the list of SemTags (or topics) and supports these initial observations with some further evidence.

Table 3 Key SemTag List: Iteration I

O _a	% _a	O _{na}	% _{na}	LL	USAS NAME
52	2.22	17	0.24	76.64	Strong Obligation
16	0.68	9	0.13	16.79	Giving
46	1.96	71	1.01	11.45	Getting & Possession
7	0.3	2	0.03	10.99	Evaluation: Good
41	1.75	65	0.93	9.5	Likely
19	0.81	21	0.3	9.34	Putting, Pulling.
17	0.73	18	0.26	8.94	Law & Order
14	0.6	13	0.19	8.85	Allowed
12	0.51	10	0.14	8.66	Employment
32	1.37	48	0.69	8.57	Politics

It shows, for example, that topics such as 'Politics' (extracted from Labour, Tories, parties, democracy), 'Law and Order' (e.g. from prisoners, jail, judges) and 'Employment' (e.g. job, work) appear in actionable comments significantly more frequently than in non actionable comments. For example, the tweet 'Get rid of the Tories all they do is cut cut cut cut cut' contains the word 'Tories' (semantic category 'Politics') and the suggestion (actionable data) of removing the Conservatives from Government.

Table 4 Key SemTags List: Iteration II

O _{na}	% _{na}	O _a	% _a	LL	USAS NAME
80	1.14	7	0.3	16.85	Numbers
306	4.37	64	2.74	12.91	Existing
259	3.7	56	2.39	8.76	Personal names
15	0.21	0	0	8.66	Evaluation: True
13	0.19	0	0	7.5	Eval: Authentic

Table 5 Key POS Lists: Iteration I and II

ITERATION I			ITERATION II		
POS	+/	LL	POS	+/	LL
VM	+	43.35	VVZ	+	12.51
VVI	+	30.39	JJ	+	11.34
VV0	+	27.29	MC	+	9.88
VM21	+	8.3	VBZ	+	9.6
TO	+	7.9	DDQ	+	7.74
POS description			POS description		
VM	modal auxiliary		JJ	General adjective	
VVI	infinitive		VVZ	"s" form of verb	
VV0	Verb base form		MC	Cardinal number	
VM21	modal auxiliary		VBZ	is	
TO	Inf.tive marker		DDQ	wh determine	

Finally, there is an overuse of the theme 'Evaluation: good' extracted from words such as 'better', which conveys the need for improvement: 'Exactly. Re-offending is the real issue here. They need better support, not better

conditions'. In the second iteration, we compare the two sets by using Set_{na} as normative text for Set_a . Table 4 shows an overuse of 'Personal Names' in non actionable tweets: Twitter users seems to share subjective opinions of no actionable 'substance' when commenting about panellists or members of the audience, as shown in the tweet '*In conclusion on: Mehdi Hasan BAD Jaqui Smith and Ming Campell TOLERABLE Francis Maude and Douglas Murray GOOD*'. The POS frequency list in Table 5 supports these initial observations.

Conclusions

This paper makes the following four key contributions: first (1), it introduces the concept of actionability as a complementary feature to automated content analysis which traditionally focuses on sentiment and topic extraction. Second (2), it applies the concept of actionability to social media analysis and finds that people spontaneously share actionable knowledge in Twitter irrespectively to the type of conversation (e.g. closed or open tweets). Third (3), it uses BBCQT as case study to investigate Twitter users' response to potential actionability 'triggers' (e.g. debates on current affairs). It finds that actionability can vary across the BBCQT data and suggests that such variation may be related to a number of factors such as the topic of discussion or the panellists on show. The automated comparative analysis between actionable and non actionable tweets (4) supports these two observations: we find an overuse of modal auxiliaries and infinitive verbs (POS) as well as topics (or semantic categories), such as 'politics' and 'employment' in actionable tweets, whereas named entities (e.g. personal names) are more frequent in not-actionable tweets. In summary, we argue that research on actionability provides a new perspective to opinion mining in that it taps into aspects of the idea generation process currently missed by traditional content analysis approaches.

Acknowledgements

The authors acknowledge support by The Research Councils UK Digital Economy Programme (grant number EP/H007237/1); The EPSRC Global Uncertainties Programme (EP/I005765/1) and EPSRC Cross-Disciplinary Feasibility Account Programme (EP/I016546/1).

References

Agichtein, E., Carlos, C., Donato, D., Gionis, A. and Mishne, G. 2008. Finding high quality content in social media. In *Proceedings of The International Conference on Web Search and Web Data Mining (WSDM '08)*. ACM, New York, NY, USA, 2008, 183 194.

Bolden, G.B. 2009. Implementing Incipient Actions: The Discourse Marker 'so' in English Conversation. *Journal of Pragmatics*. Vol. 41, Issue 5, pp. 974 998.

Boyd, D., Golder, S. , Lotan, G. 2010. Tweet, Tweet and Retweet: Conversational Aspects of Retweeting on Twitter. In *Proceedings of 43rd Hawaii International Conference on System Sciences* (2010) pp. 1 10.

Cao, L. and Zhang, C. 2006. Domain Driven Actionable Knowledge Discovery in the Real World. *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 821 830.

Cataldi, M., Di Caro, L. and Schifanella, C. 2010. Emerging Topic Detection on Twitter Based on Temporal and Social Terms Evaluation. In *Proceedings of the Tenth International Workshop on Multimedia Data Mining (MDMKDD '10)*. ACM, New York, NY, USA, Art. 4.

Ghosh, A. and McAfee, P. 2011. Incentivizing High Quality User Generated Content. In *Proceedings WWW 2011*, March 28 April 1, 2011, Hyderabad, India. ACM 978 1 4503 0632

Honeycutt, C. and Herring, S.C. 2009. Beyond Microblogging: Conversation and Collaboration via Twitter. In *Proceedings of 42nd Hawaii International Conference on System Sciences*. (2009) pp. 1 10.

Hopkins D. and King, G. 2010. A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science*, 54 (1). 229 247.

Java, A., Xiaodan, S., Finin, T. and Tseng, B. 2007. Why we Twitter: Understanding Microblogging Usage and Communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis (WebKDD/SNA KDD '07)*. ACM, New York, NY, USA, 56 65.

Neubig, G., Matsubayashi, Y., Hagiwara, M., & Murakami, K. 2011, Safety Information Mining What Can NLP Do in a Disaster. *Proc of 5th Intl. Joint Conf. on Natural Language Processing*, Chiang Mai, Thailand, (2011), 965 973.

Pang, B. and Lee, L. 2008. Opinion Mining and Sentiment Analysis. *Found. Trends Information Retrieval*. 2, 1 2 (January 2008), 1 135.

Rayson, P. 2008. From Key Words to Key Semantic Domains. *International Journal of Corpus Linguistics*, 2008, 13 (4). pp. 519 549. ISSN 13846655.

Searle, J.R. 1969. *Speech Acts Cambridge University Press*, ISBN 0 521 09626 X.

Simm, W.A., Ferrario, M.A., Piao, S., Whittle, J. and Rayson, P. 2010. Classification of Short Text Comments by Sentiment and Actionability for VoiceYourView. In *Proceedings of the Second IEEE International Conference on Social Computing (SocialCom2010)*, Minneapolis, MN, USA, 2010, 552 557.

Thomas, J. 1995. *Meaning in Interaction: An Introduction to Pragmatics*, Longman. ISBN 10: 9780582291515

Whittle, J., Simm, W.A., Ferrario, M.A., Frankova, K., Garton, L., Woodcock, A., Nasa, B., Binner, J. M. and Ariyatun, L. 2010. VoiceYourView: Collecting Real time Feedback on the Design of Public Spaces. In *Proceedings of the Ubicomp Conference 2010*, Copenhagen, Denmark, 2010, pp. 41 50.

Xiaofeng, H. Ding, C.H.Q., Zha, H., Simon, H.D. 2001 Automatic Topic Identification Using Webpage Clustering. In *Proceedings IEEE International Conference on Data Mining*, 2001. ICDM 2001, pp.195 202.

Zhao, D. and Rosson, M.B. 2009. How and why people Twitter: the role that micro blogging plays in informal communication at work. In *Proceedings of the Int. Conference on Supporting Group Work (GROUP '09)*. ACM, New York, NY, USA, 24.