Knowledge Discovery from Social Media Data:
A Case of Public Twitter Data for SMEs

Christopher Adetunji, Leslie Carr
Web Science Institute
School of Electronics and Computer Science
University of Southampton
Southampton, England
Email: {ca6g14, lac}@soton.ac.uk

Abstract—**Making sense of social media data is increasingly becoming a subject of concern to corporate organisations. It is therefore, no coincidence that the subject of Knowledge Identification and Discovery is currently receiving a huge attention within industry and academia. Research has shown that there is an enormous wealth of actionable knowledge to be gained from social media data for organisations' strategic competitive advantage. However, this opportunity is not being harnessed by Small and Medium-sized Enterprises (SMEs) as much as it is by larger enterprises. This is due, in part, to a misconception that social media is not that relevant to SMEs as much as it is to larger corporations. This paper presents a qualitative exploratory study, which attempts to show that social media can be mined for organisational knowledge that is relevant to the strategic competitive advantage of SMEs. A case of a medium-sized enterprise, which is previously without a significant social media presence, is explored with regards to how public Twitter data is exploited to discover actionable knowledge that propels the enterprise's strategic competitive advantage.**

Keywords—social-media; data; twitter; SME; knowledge.

I. INTRODUCTION

Beyond their use for building relationships, connections, and/or marketing leads, social media present an opportunity for Small and Medium-sized Enterprises (SMEs) — just as they do for large corporate enterprises — to exploit the wealth of intrinsic insights embedded in the mass of social data publicly available, for their strategic competitive advantage. Yet, due to a perceived lack of relevance of social media to certain types of industry/sector, SMEs are often disinclined to adopting social media [1]. This work presents a case of a medium-sized enterprise for which social media is perceived as not relevant. Since the organisation is without a huge social media presence, this work utilises public Twitter data, which are mined for relevant insights, and the knowledge gained are actioned for strategic competitive advantage. The pieces of knowledge extracted are explored with a discussion on their actionability as well as other valuable insights potentially embedded in such public social data.

Background

Social media has been adopted by organisations to support both the individual and corporate Knowledge Management processes [2]. It forms a social machine that facilitates human interactions on the Web [3], enabling people to create new knowledge by sharing and synthesising knowledge from various sources. This is aided by the technological platforms upon which social media tools are built, which facilitate the cognitive processes previously performed by people [4][5]. Essentially, *social media* includes tools and software platforms that enable humans to participate in the social process of content and knowledge generation, collaboration and knowledge sharing. Social media trends began with the rise of so-called Web 2.0 [6], in which sites became sophisticated apps and content-management platforms designed to facilitate the creation and sharing of user-generated data and content [7]. These platforms include social sharing and networking tools like Facebook, Twitter, blogs, wikis and forums [8][9][10]. In addition, [3, p.2] identifies mySpace, Ushahidi, Galaxy Zoo, reCaptcha and Wikipedia as social software exemplars of social machines.

With their support for contributions and knowledge sharing from across a wide range of avenues (e.g., tweeting via Short Messaging Services (SMS) or smart phone apps), social media tools enhance knowledge sharing within the organisation in their capacity for fostering discussions over documents and thereby enabling organisations to build social environment or communities of practice necessary for facilitating the sharing of tacit knowledge [11][12, p.26]. This has impacted the strictly-controlled world of corporate Information Technology (IT) services, creating an agenda of Enterprise Mobility that is implemented by employee-owned devices adapted for company use, a concept commonly referred to as Bring Your Own Device (BYOD); and/or company-owned devices that support personal use, also commonly referred to as Company-Owned Personally-Enabled (COPE). Consequently, an enormous amount of rapid and varied data is being produced by social applications embedded in the workplace, potentially available for organisations' insight using techniques of big data analytics such as text analysis of unstructured data sources and massively parallel processing (MPP) to analyse streaming data for informed decision and better results [13][12]. This work utilises text analysis techniques to make sense of the unstructured social media data harvested through Twitter's Streaming API.

As a social micro-blogging utility, Twitter creates increased interest in organisations with regards to growth, features and potential benefits to the organisation [14]. For example, apart from its significant role in the US Elections of 2008 [15], [6] also alluded to Dell's claim that its use of Twitter has generated $1 million in incremental revenue due to sales alerts. Meanwhile, [16] describes Twitter as 'a glorified piece of valuable infrastructure that enables rapid and easy communication' and, unlike Facebook or LinkedIn, its asymmetric

relationship model of 'following' allows one to keep up with the tweets of any other user without the need for the other user to reciprocate. This facilitates a lateral flow of knowledge that is powered by the intrinsic motivation of individual employees within the organisation. Moreover, the consumerisation and proliferation of mobile devices like smart phones has driven the popularity of social media and enabled the adoption of Web 2.0 affordances, especially the deployment of micro-blogging, to the business environment. This offers powerful opportunities to distribute 'tacit knowledge' and 'best practices' within an enterprise [15].

An increasing number of large enterprises have already been able to tap into the benefits of Twitter as a micro-blogging platform. According to a Gartner report referenced in [15], leading-edge companies are investigating the potential of micro-blogging to enhance other social media and channels, and, as mentioned above, Dell recounts its use of Twitter as a leverage for increased revenue gains while the electoral success of Barack Obama as president in the US General elections of 2008 is largely credited to the use of Twitter by the Democratic Party [15]. Also, Ford Motors and Zappos [17] are a few examples of large enterprises already exploiting social media. How could this trend be beneficial to Small and Medium sized enterprises even where the social applications may not be hugely embedded in the work place? This work presents how a medium-sized enterprise was able to tap into the wealth of Twitter data for its operational and strategic insights.

II.   CASE STUDY METHODOLOGY

The role of Small and Medium-sized Enterprises within an Economy is so crucial that the World Bank commits hugely to the development of the sector as a significant part of its efforts in promoting employment and economic growth [18]. *Liaise Loddon* is a medium-sized enterprise with about 220 employees, providing residential social care for adults with autism and learning disabilities in Hampshire, United King-dom. As typical in this sector, operational procedures result in an enormous amount of documentation arising from daily diaries, incident/activity reports and several other reporting in compliance with regulatory requirements, analytical purposes and decision making. Although the company has recently deployed an enterprise mobility suite of mobile devices and applications to replace the existing paper-based documentation system, this experiment explores how this enterprise mobil-ity agenda could be hardened with knowledge sharing and knowledge extraction from the mass of social data freely available on Twitter, for example, in such a way as it supports the organisation at the second level of organisational change, which highlights the people dimension of a socio-technical system [19, p.35-38]. This work utilises simple textual analysis techniques to make sense of the unstructured social media data harvested through Twitter's Streaming API. This is a practical approach that is replicable with a cost of next to nothing.

Ordinarily, Liaise Loddon's operations do not require social media marketing, neither does it appear like the company could benefit from its employees' conversations and knowledge sharing over social blogging platform like Twitter. As such, this organisation, just like many other small and medium-sized enterprises (SMEs), does not have a huge following

on Twitter, neither does it have any such enterprise micro-blogging platform that generates sufficient data from which employees' conversations could be mined for insights. Yet, in a bid to stay abreast of — and respond quickly to — issues surrounding its area of specialism, this work harvested, from global public tweets, for *Autism*, and its variants like 'ASD' and 'disability', which are categorical keywords that tend to define aspects of its business domain.

III.   DATA GATHERING AND FILTERING

For the experiment, a week's worth of tweets are harvested (from 30th April to 6th May, 2015) with a toital of 149,501 from the Twitter streaming API and, using textual analysis technique, extraneous elements are filtered out in order to reduce the data. In data mining, one solution to the challenges of handling vast volumes of data is to reduce the data for mining [20].

TABLE I: 3.5% OF TWEETS WITH LOCATION ENABLED

		Percentage (%)
No. of Tweets without Country Value	144246	96.5
No. Tweets with Country Value	5255	**3.5**
Total No. of Tweets	**149501**	100

This dataset is a microcosm of the global public tweets that feature the categorical keywords of interest as mentioned above. In reducing the data, we also need to remove tweets that are not in English language. For example, the French word 'autisme', which is the same as 'autism' in English, may have had the letter 'e' mistakenly omitted by a person tweeting in French and this would have resulted in the French tweet being harvested. As a way of delimiting the scope of the study, tweets that are not in English language are filtered out. Also, the abbreviation, 'ASD', would not always represent our intended use neither would it always represent *Autistic Syndrome Disorder* even in the English language.
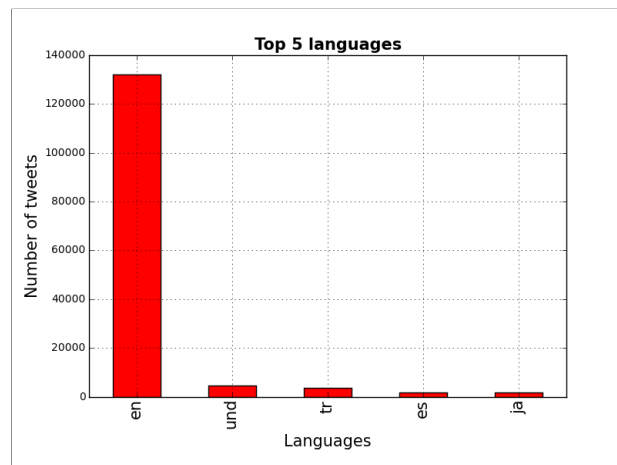

Figure. 1: The Top 5 Languages of Tweets Polled

In essence, those tweets containing the polled keywords are filtered out where they are deemed to be out of context. Moreover, we could only determine the country for about 3.5% of tweets in the harvested Twitter data, as indicated in Table I. This is due to Twitter users not updating their profile with
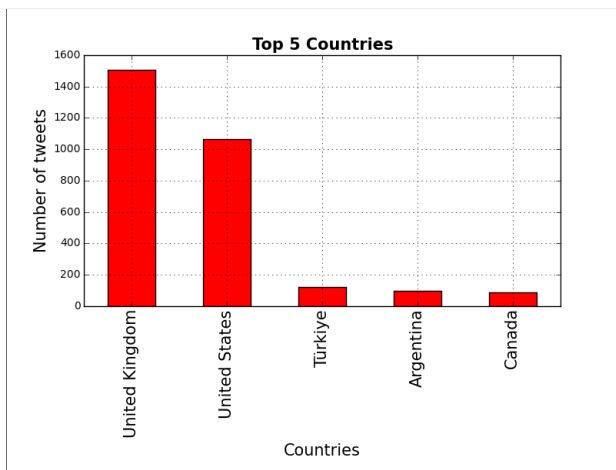
Figure. 2: The Top 5 Countries of Tweets Polled

their locations as well as some imperfection with Twitter's geolocation algorithm [21].

Therefore, to begin making sense of this data, we concentrate on the top 5 languages and top 5 countries as presented in Figures 1 and 2 respectively, out of which the experiment was narrowed to tweets in English language and from the United Kingdom. *(Please note the following expanded abbreviations as used in Figure 1: en, English; und, Undecided; tr, Traditional Chinese; es, Spain; ja, Japan.)*

TABLE II: CONTENT CLASSIFICATION OF TWEET DATA

Contents	No. of Tweets (Including RTs)
Impact of Technology on Disability	15
Information Gathering	10
Political Opinions (#votecameronout)	132
Social Welfare Benefits	327
Living with Autism	989
Total Tweets	**1473**

As indicated in Figure 2, there are 1,473 tweets emanating from the United Kingdom. Using regular expressions, the tweets are classified according to the contents as indicated in Table II.

IV. FINDINGS AND KNOWLEDGE EXTRACTED FROM THE DATA

Despite the data collection being based on domain-specific keywords of interest to this paper's case study, the research is an exploratory study in which there was not a preconceived idea of the insights/knowledge inherent in the data. Out of an enormous amount of data, only a handful may contain the valuable and actionable knowledge that propels an organisation towards strategic competitive advantage [20, p.5]. As such, the bulk of the contents as seen in Table II, are largely re-tweets (RT) of the original messages and so, may be regarded as extraneous amplification of the original tweets. Therefore, this section describes the categories observed in the data and the next section follows with a discussion on the value and actionability of the knowledge so discovered:

1) **Impact of Technology on Disability**

  "*RT @BILD_tweets: Helping to unlock the secrets of autism - a project using innovative technology aims to change how we address autism http:...*"

The above tweet provides an insight into a project using innovative technology to change how we address *autism*. As this paper's case study organisation is in the business of autism support and also currently implementing mobile technologies to enhance its operational performance, it is worth exploring this piece of insight further.
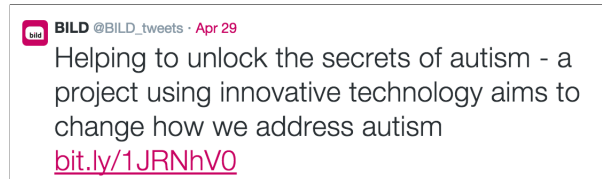

Figure. 3: Original Tweet with Link to Project on Innovative Technology

Although the link to the actual URL of the story about the project is missing from the tweet, we can easily follow up with the original source of the tweet, as the above is a RT (Re-Tweet) of @BILD_tweets, which is the Twitter handle for BILD (British Institute of Learning Disabilities). BILD actually tweeted that piece of content on the 29th of April, which is a day before our data capture began, as can be seen in Figure 3. This explains why the original tweet was not captured in our twitter streaming data capture of 30th April to 6th May. From this original tweet, we have been able to extract the URL link (bit.ly/1JRNhV0) to the story about the project on innovative technology. This is about the National Autism Project, which "aims to create a more strategic approach to addressing the challenges of the condition". This project highlights the impact of iPads, picture dictionaries and interactive schedules on the improvements of communication and vocabulary of autistic people. Strategic competitive advantage requires an alignment/tagging along with this project. Below are samples of other tweets related to this content of Technology's impact on disability while its pertinence, as an actionable piece of knowledge, is discussed further in Section V:

  "*tech reducing the impact of disability - or are the latest gadgets too pricey? Watch @SkyNewsSwipe at 2130 http://t.co/iHtX1spOqQ*"

  "*Technology limits impact of disability but is it affordable? @TwitterUser_GT http://t.co/Az3nJejO32*"

2) Information Gathering

"@TwitterUser @BBCNewsUS @BBCWorld Please direct me to this research, the thing about vaccines causing autism was admitted to be a fraud."

The first tweet about vaccines causing autism in this category is a request for information. Just as an enterprise micro-blogging tool could be used within the organisational social network, public micro-blogging tools like Twitter provide the platform to quickly seek information, knowledge and/or ideas from a heterogeneous audience defying the constraints of space, time and location. Thus, the above tweet was almost instantly replied to by the one below:

"@TwitterUser Here's the original study that said that vaccines cause autism, from a respected, peer-reviewed journal: http://t.co/cmVVKpLQgh"

Even though the original study is from a 'respected, peer-reviewed journal', as claimed by the sender of the above tweet, we know from the link provided that the publication of the research has been retracted as shown in Figure 6. The ability for anyone to search, gather and distribute information seamlessly in this manner provides an interesting dimension of social media as "relatively inexpensive and widely accessible electronic tools that enable anyone to publish and access information..."[22].
Meanwhile, the following two tweets provide link to further information that could help drive home the knowledge that the research study in question has actually been rebuffed:

"RT @TwitterUser: @SB277FF vaccines do not cause autism. They don't. But if they did, what would you prefer? Autism or incurable smallpox/po"

""RT @BILD_tweets: There is 'no link between MMR and autism', major study concludes. http://t.co/Re9L8fPfGV via the @guardian #autism""

In as much as Twitter allows for an almost spontaneous expression of opinions by anyone, it offers a good platform for healthy debate on topical issues from which knowledge could be mined, as exemplified by the question of preference between autism and incurable smallpox posed by one of the tweets above.

Moreover, the following tweet with a URL link to *Learning Disability Census* is an example in knowledge discovery (of an official census and regional data on Learning Disabilities), which when actioned in conjunction with the enterprise resource planning, could have an impact on the company's strategic planning:

"RT @dmarsden49: Learning disability census with regional stats is out. Check http://t.co/Ja3tk7ZRDZ"

Figure. 4: The Proliferated Re-tweets of Political Opinion

3) Political Opinions (#votecameronout)

The role of public opinion cannot be over-emphasised insofar as it shapes and is shaped by government policies. A recent and relevant example is the UK tax credits row [23], which has seen the planned tax credit cuts, at the time of writing this report, suspended by government because the scheme proved unpopular to the public and thus defeated in the House of Commons. Social media, especially Twitter, provides a means of capturing and measuring the sentiments and opinions of the electorate.
It is therefore, no coincidence that political opinions that have been expressed, are included in the Twitter data gathered over *autism* and *disability* keywords:

"#votecameronout Because he wants to get rid of Human Rights Act which will affect: Maternity Rights; Workers rights; Disability Rights"

"For the harassment of people struggling on sick & disability benefits... #VoteCameronOut"

"5 more years of the Tories we will lose Social Care, NHS, Human Rights, Workers Rights, Unions, Disability support. #VoteCameronOut"

Using the hashtag #votecameronout in the run up to the UK General Elections of 2015, the above tweets represent an active campaign against the then incumbent Tory-led government in which David Cameron is Prime Minister. It is interesting to note that the bulk (129) of the political tweets in this experiment's Twitter data are a proliferated re-tweets (RT) of the above 3 original tweets as exemplified in Figure 4. The correlation between public sentiments on social media and elections results and/or on government policies, is another growing area of interest in social media research. In politics meanwhile, it is not uncommon for opponents to whip up public sentiments by whatever means possible. Social Welfare issues are quintessentially core, and often politicised, concerns in the UK. A parallel category of tweets in this work is that of social welfare benefits, which is described in the next section. Although this research's data-set is based, as stated earlier, on categorical keywords that define the business of the case study organisation, the infiltrated political opinions cannot be ignored in as much as these are public opinions that shape political trends which potentially impacts on businesses in terms of government policies. Akin to this is the category on
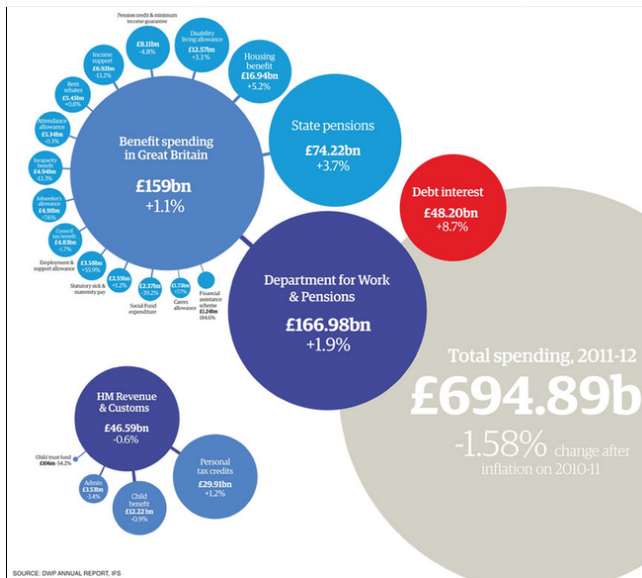
Figure. 5: Public Spending on Benefits in the UK

social welfare benefits, described in the next section.

4) **Social Welfare Benefits**

Social Welfare simply implies the "Well being of the entire society" [24], which promotes inclusivity for the disabled, the sick, the elderly/pensioner, the unemployed and even the low income earners. As this is the hallmark of an egalitarian society, the UK government renders financial assistance to these categories of people in form of a range of Social Welfare Benefit payments. Figure 5 provides an insight into public spending on social welfare benefits in the UK. [25]. As indicated in the preceding section, social welfare issues affect the fabrics of the society and any proposed significant cut in social welfare benefits is a natural invitation for public dissent. This category of tweets from this work represents genuine sentiments and opinion of those expressing them, without political motivations like the preceding category:

"*Uproar at thought of @Conservatives cutting child benefits if elected - I wish there was same media outrage over disability cuts #GE2015*"

"*@George_Osborne If only I could live until pensionable age. You've reduced my disability benefit well below living standards!*"

"*39 yo woman killed herself after Department Work and Pensions threats to cut off disability benefits http://t.co/TkVQF2UYki...*"

Again, the above are a few samples of sentiments and opinions about *Child Benefits* and *Disability Benefits*, which provide an initial understanding to the unassuming, that social welfare benefits are not a one-size-fits-all affair but are multifarious (see Figure 5), with some being exclusively non- means tested

(e.g, Child Benefit). These tweets provide some insights into public sentiments towards government policies. Since any of such social welfare benefit cuts would directly and/or indirectly impact the service users and providers of social care, it can be inferred that the case study organisation would also share these public sentiments.

5) **Living with Autism**

Autism is defined as a *life-long neurodevelopmental condition interfering with the person's ability to communicate and relate to others* [26]. How can this definition be juxtaposed with one of the myths surrounding autism [27, item 8] that *autistic people do not interact*? This myth is however, dispelled by the tweet below, which is a re-tweet of an original tweet by an actual autistic blogger who attempts to use his blog posts to connect with the general public:

"*@matt_diangelo RT? It would be truly amazing if u could view my blog about living with Autism&OCD. Would mean a lot- http://t.co/JCGBBZz8fJ*"

This category constitutes the bulk of the Twitter data for this work as it contains multiple unique re-tweet of the same tweet over 900 times (see Table II. This is an indication of the public interest/curiosity and positive sentiment towards the subject of autism in general, and towards the autistic blogger in particular. Despite the National Health Service (NHS)'s attempts at educating the general public by diffusing some of the myth surrounding the subject of *Autism* [28], among several *Autism Awareness* initiatives, the story of autism as told by an autistic person appears to garner more public support and understanding .

Measuring public opinion and sentiments through social media impact, reach and networks is another interesting research area in social media research towards which this work can potentially be extended.

V. ACTIONABILITY OF KNOWLEDGE EXTRACTED

The real essence of knowledge is its actionability, especially when it contributes to the advancement of a proposed undertaking [29]. Each of the knowledge items discovered from the tweet data, as highlighted above, is capable of providing significant insights that would inform decision making, which impacts company's proposed undertakings at one point or the other. However, the first item, *Impact of Technology on Disability*, (No.1) is more pertinent to the enterprise mobility agenda by which the company deploys mobile application and devices to its operations. For example, one of the shortened URL above (http://t.co/Az3nJejO32) leads us to a Sky News supplement on *'How Tech is Helping with Struggle of Disability'*. To aspire to a leadership position in the health and social care sector, the case study organisation cannot afford to be oblivious to such reports as this, which could potentially shape the industry trends and direction. This knowledge, coupled with the insights gained from the use of iPads and pictorial dictionary mentioned in the *'Project on Innovative Technology'* resulted in an official resolution by the company to extend the use of mobile devices to its service users as well, and not only

to help staff in operational performances. It is worth noting that, although the piece of actionable knowledge regarding the impact of technology on disability was on the news prior to the extraction of data for this work, it was neither known nor acted upon by the company until the above decision was driven through a presentation made by the authors of this paper.

Meanwhile, without a preconceived agenda, this work has been able to discover actionable knowledge and strategic insights from Twitter. As discussed, these are, that the use of technology being embarked upon by the organisation is not misguided; that the widespread belief that MMR vaccines cause autism has been debunked; and the regional distribution of autism through the *Learning Disability Census* about autism; among others.

Apart from the above insights, there are more valuable insights that an organisation could derive from these kinds of public social data. These include:

1) *Expertise location:*
The micro-blogging, hash-tagging and re-tweeting affordances of Twitter places its contents in the public space, as opposed to emails or local files, which are prone to privacy concerns. According to [10], "the diversity of both the content type and the user associations with contents is an indication that expertise information derived from social media data can be of great value". [10] further assert that social media that reside behind a firewall (e.g., *Yammer* — a corporate Twitter clone [30]), is typically used by employees to discuss internal topics, and hence reflects the organisation's unique vocabulary and areas of interest. This enables the organisation to find people (or employees) who are knowledgeable in a given topic.

2) *Public Opinion and Sentiments:*
Public opinion about the organisation's image, product or services can make the difference between success and failure for the organisation. Knowing the sentiments of the public would help the organisation in responding in such a way that would ginger or maintain favourable public opinion. Businesses have made efforts to find out customers' sentiments and opinions, often expressed in free text, towards company's products and services [31]. Twitter enables expressions in free text, which is why the bulk of Twitter data is textual. [12] cited a real world example of a client who introduced a different kind of environmentally friendly packaging for one of its staple brands. Customer sentiment was somewhat negative to the new packaging, and some months later, after tracking customer's feedback and comments, the company discovered an unnerving amount of discontent around the change and therefore moved to a different kind of eco-friendly packaging. They therefore, hypothesise that, "*if you dont have some kind of micro-blog oriented customer sentiment pulse-taking going on at your company, you're likely losing customers to another company that does.*"
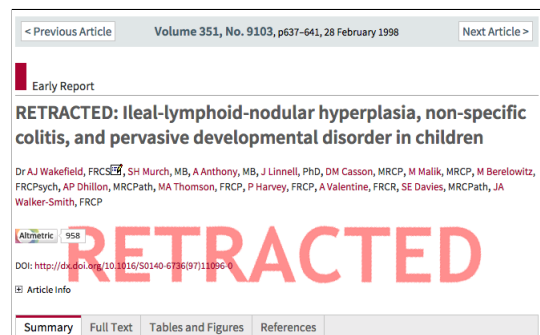

Figure. 6: Retracted Study Linking Vaccine to Autism

VI. CONCLUSION

This paper has attempted to explore the viability of social media as rich data sources from which SMEs — like large enterprises — can discover knowledge and insights for their strategic competitive advantage. The paper presented Twitter as a prolific data source of all social media and examined the case of a medium-sized enterprise for which public Twitter data was explored and exploited for valuable insights without a preconceived idea of the insights/knowledge inherent in the data. The paper further examined the contents of the social data with regards to the actionability of the pieces of knowledge so discovered, which helped in demonstrating that out of an enormous amount of data, only a handful may contain the valuable and actionable knowledge that propels an organisation towards strategic competitive advantage. In view of this, this paper posits that small and medium sized enterprises are not precluded from using social media data to augment their corporate knowledge assets through knowledge discovery, whether or not they have a huge presence or following on social media. The paper concludes therefore, that any enterprise of any size can explore and exploit public social data to its organisational advantage whether or not they are players in the social media sphere. Meanwhile, the structure of the connections and relationships within a social network like Twitter can be visualised to provide further depth and insights to the pieces of knowledge discovered from the network. This work can be extended to cover these terrains in future studies.

REFERENCES

[1] N. Michaelidou, N. T. Siamagka, and G. Christodoulides, "Usage, barriers and measurement of social media marketing: An exploratory investigation of small and medium B2B brands," *Industrial Marketing Management*, vol. 40, no. 7, Oct. 2011, pp. 1153–1159. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0019850111001374 [Last Accessed 01 March 2016].

[2] L. Razmerita, K. Kirchner, and T. Nabeth, "Social Media in Organizations: Leveraging Personal and Collective Knowledge Processes," *Journal of Organizational Computing and Electronic Commerce*, vol. 24, no. 1, Jan. 2014, pp. 74–93.

[3] P. R. Smart and N. R. Shadbolt, "Social Machines," in *Encyclopedia of Information Science and Technology*, M. Khosrow-Pour, Ed. Pennsylvania: IGI Global, Aug. 2014.

[4] M. Sigala and K. Chalkiti, "Knowledge management, social media and employee creativity," *International Journal of Hospitality Management*, vol. 45, 2015, pp. 44–58. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S027843191400173X [Last Accessed 02 March 2016].

[5] L. Carr and S. Harnad, "Offloading cognition onto the Web," *IEEE Intelligent Systems*, Jan 2011. [Online]. Available: http:

//eprints.soton.ac.uk/271030/1/lesdoc2.pdf [Last Accessed 02 March 2016].

[6] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," *Business Horizons*, vol. 53, no. 1, 2010, pp. 59–68.

[7] M. Kleek and K. O'Hara, "The Future of Social is Personal: The Potential of the Personal Data Store," 2014. [Online]. Available: http://eprints.soton.ac.uk/363518/1/pds.pdf [Last Accessed 02 March 2016].

[8] D. P. Ford and R. M. Mason, "A Multilevel Perspective of Tensions Between Knowledge Management and Social Media," *Journal of Organizational Computing and Electronic Commerce*, vol. 23, no. 1-2, Jan. 2013, pp. 7–33.

[9] A. P. McAfee, "Enterprise 2 . 0 : The Dawn of Emergent Collaboration," *MIT Sloan Management Review*, vol. 47, no. 3, 2006, pp. 21–28.

[10] I. Guy, U. Avraham, D. Carmel, S. Ur, M. Jacovi, and I. Ronen, "Mining expertise and interests from social media," *Proceedings of the 22nd international conference on World Wide Web*, 2013, pp. 515–526.

[11] T. H. Davenport and L. Prusak, *Working Knowledge: How Organizations Manage what They Know, Part 247*. Boston: Harvard Business Press, 1998.

[12] P. Zikopoulos, C. Eaton, D. DeRoos, T. Deutsch, and G. Lapis, *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*, S. Sit, Ed. New York: McGraw Hill, 2012.

[13] J. Lamont, "Big data has big implications for knowledge management," *KMWorld*, vol. 21, no. 4, 2012. [Online]. Available: http://www.kmworld.com/Articles/Editorial/Features/Big-data-has-big-implications-for-knowledge-\\management-81440.aspx [Last Accessed 02 March 2016].

[14] M. Bulearca and S. Bulearca, "Twitter: a Viable Marketing Tool for SMEs?" *Global Business and Management Research: An International Journal*, vol. 2, no. 4, 2010. [Online]. Available: http://www.gbmr.ioksp.com/pdf/BulearcaandBulearca,2010.pdf [Last Accessed 02 March 2016].

[15] M. Böhringer and A. Richter, "Adopting social software to the intranet: a case study on enterprise microblogging," in *Proceedings of the 9th Mensch & …*, 2009, pp. 1–10. [Online]. Available: http://www.kooperationssysteme.de/docs/pubs/BoehringerRichter2009-M&C-Enterprise%20Microblogging.pdf [Last Accessed 02 March 2016].

[16] M. A. Russell, *Mining the Social Web, 2nd Edition - O'Reilly Media*. O'Reilly Media, 2013.

[17] A. Porterfield, "9 Companies Doing Social Media Right and Why," 2011. [Online]. Available: http://www.socialmediaexaminer.com/9-companies-doing-social-media-right-and-why/ [Last Accessed 02 March 2016].

[18] M. Ayyagari, T. Beck, and A. Demirguc-Kunt, "Small and Medium Enterprises Across the Globe," *Small Business Economics*, vol. 29, no. 4, 2007, pp. 415–434.

[19] G. Piccoli, *Information Systems for Managers: Texts & Cases*. Hoboken: John Wiley & Sons, 2008.

[20] T. W. Liao and E. Triantaphyllou, *Recent Advances in Data Mining of Enterprise Data: Algorithms and Applications*. World Scientific Publishing Company Pte Limited, 2008.

[21] M. Graham, S. A. Hale, and D. Gaffney, "Where in the World Are You? Geolocation and Language Identification in Twitter," *The Professional Geographer*, vol. 66, no. 4, May 2014, pp. 568–578.

[22] D. Murthy, *Twitter: Social Communication in the Twitter Age*. Cambridge: Polity Press, 2013.

[23] "Tax Credits Row What Will Happen Now After Government Defeat?" 2015. [Online]. Available: https://uk.news.yahoo.com/tax-credits-row-happen-now-100031983.html [Last Accessed 29 February 2016].

[24] "What is social welfare? definition and meaning." [Online]. Available: http://www.businessdictionary.com/definition/social-welfare.html [Last Accessed 03 March 2016].

[25] S. Rogers, "UK welfare spending: how much does each benefit really cost?" 2013. [Online]. Available: http://www.theguardian.com/news/datablog/2013/jan/08/uk-benefit-welfare-spending [Last Accessed 30 April 2016].

[26] M. Elsabbagh *et al.*, "Global prevalence of autism and other pervasive developmental disorders." *Autism research : official journal of the International Society for Autism Research*, vol. 5, no. 3, Jun. 2012, pp. 160–79.

[27] J. Baldock, "World Autism Awareness Day 2015: 16 autism myths debunked," Apr. 2015. [Online]. Available: http://metro.co.uk/2015/04/02/autism-sixteen-myths-debunked-5123051/ [Last Accessed 03 March 2016].

[28] N. Choices, "Autism spectrum disorder - NHS Choices." [Online]. Available: http://www.nhs.uk/Conditions/Autistic-spectrum-disorder/Pages/Introduction.aspx [Last Accessed 03 March 2016].

[29] R. Cross and L. Sproull, "More Than an Answer: Information Relationships for Actionable Knowledge," *Organization Science*, vol. 15, no. 4, Aug. 2004, pp. 446–462.

[30] J. Zhang, Y. Qu, J. Cody, and Y. Wu, "A case study of micro-blogging in the enterprise," in *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. New York, USA: ACM Press, apr 2010, p. 123.

[31] C. Lin and Y. He, "Joint sentiment/topic model for sentiment analysis," in *Proceeding of the 18th ACM conference on Information and knowledge management - CIKM '09*. New York, USA: ACM Press, Nov. 2009, p. 375.