

Social Media and Big Data – Cracks in the Crystal Ball?

Citation for published version (APA):

Westera, W. (Author). (2013). Social Media and Big Data – Cracks in the Crystal Ball?. Web publication/site, Research World.

Document status and date:

Published: 14/10/2013

Document Version:

Peer reviewed version

Document license:

CC BY-NC-ND

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

<https://www.ou.nl/taverne-agreement>

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 12 Oct. 2022

Open Universiteit
www.ou.nl



Social Media and Big Data – Cracks in the Crystal Ball?

<http://rwconnect.esomar.org/using-social-media-for-market-analysis-cracks-in-the-alleged-crystal-ball/>

Wim Westera is a full professor of digital media at the Open University of the Netherlands. He works at the crossroads of social sciences, cultural sciences and computer science. He holds a PhD in physics and mathematics at Utrecht University. After a stay with the BBC he has made dozens of films and television programmes. At the University he works as a digital media specialist targeting educative simulations, multimedia, e-learning and serious gaming. He is the author of “The Digital Turn” about the impact of digital media on our existence (www.thedigitalturn.co.uk).

New ways of market research

Similar to crystal-gazers, palmists and dowzers today’s market researchers are driven by the ambition to foretell the future. Their profession is to cautiously trace and gather data of businesses and markets and anticipate future developments. Just like old times it is safe to say that good predictions make good money. Today’s breathtaking time of emerging digital media and the ongoing virtualisation of the world are bringing about fundamental changes in the way our society functions. The digital products and services of IT companies such as Google, Microsoft, Apple, Facebook and Twitter have pervaded our daily lives and have induced new modes of communication and interaction. In less than a decade the internet, smartphones, tablets, GPS and social media have radically changed the functioning of business, economy, social life and culture as a whole. Presently, market researchers have widely adopted social media as an effective and affordable tool for gaining insights into target groups, customer needs, brand impact and other important market research. Collecting and analyzing the ongoing data stream of social media hold the promise of gaining more insights, faster insights and better insights about future events and developments. We’ve replaced the crystal ball with social networks.

The social media wave

Today, over five billion people have mobile phones, two billion people are on the internet, and Facebook is the world’s largest community serving over one billion of users (they say). Now that the individual spends on average more than 3 hours online, social media have become so dominant that they cannot be ignored. In social media each user actively produces data, either by entering keywords in a search engine, by sharing tweets, comments, profiles, favorites, likes or follows, or by uploading or downloading blogs, stills, videos or any other content. Every single action produces added value, which can be aggregated and capitalized in commercial advertising and market analysis. This is how companies such as Google and Facebook have been able to become multibillion dollar businesses. The many-to-many nature of social networking produces data that directly reflect what users actually do and what they’re talking about. Now we’ve got rid of the drawbacks of traditional methods, such as questionnaires, polls, interviews and panels, which often take opinions for facts and fail to compensate for the differences between what people say and what they actually do. By now, we’re unmistakably living in the era of big data. The amount of information on the internet is inconceivably large: it doubles every 18 months. According to basic principles of statistics having more data means achieving higher predictive powers. Indeed there are quite some stunning examples. A study of the key words that people entered in Yahoo’s and Google’s search engines has

shown that the development of flu epidemics can be now identified much earlier than by the official virus surveillance methods ([Nature](#)). A recent Google study has shown that the search query volume of moviegoers is an accurate predictor of movie box office performance ([Google Whitepaper](#)). Unmistakably, internet datamining has great potential because it entails short-cycled, continuous monitoring, observable behaviours rather than opinions, and big data aggregates with enormous predictive power. For market analysis social media seem to get close to a goldmine.

Some sneaky issues

Unfortunately this isn't all true. Using social media may be a promising approach, but having achieved some successes doesn't necessarily prove the method's validity: crystal gazers were occasionally right despite dubious methods. Social media are a relatively new platform for human communication. In my book "The Digital Turn" I explain how the invention of new media (e.g. clay tablets, the alphabet, print, computers) has allowed us to sustainably express, consolidate and share knowledge and thereby enabled us to develop our culture and society. The book also anticipates the impact of recent and future developments such as social media, ubiquitous computing, unlimited storage capacity, big data, mixed and augmented realities and super-human intelligence. It analyses how our digitally-enhanced biotope alters our social interactions, our behaviours, and ultimately the ways we arrange our lives, and makes suggestions how to cope with current and future media technologies. Quite some issues covered by the book are applicable to internet-based market analysis. Let look at some of the flaws of social media analytics.

Biased samples

First, getting your data from services such as Facebook or Twitter will burden you with highly biased samples with an overrepresentation of persons in their twenties or thirties, and highly involved in digital media usage. Altogether this isn't very different from the bias in election polls, which are likely to attract mainly persons that are interested in politics in their samples. Strikingly, screaming headlines appear when a poll indicates a shift of one seat, thereby neglecting the poll's poor predictive power, which may easily be ten seats up or down. Likewise, results of social media are often dubious.

Are we basing on quicksand?

Secondly, we should realize that by listening to social conversations we base ourselves on what we may fairly qualify as smalltalk and gossip. Although evolutionary biologists claim that gossip is an intrinsic human function that is directly linked with the preservation and amplification of solidarity in groups and the establishment of community codes and norms, its content is readily associated with scandal, rumours and misinformation. Hence, tapping into the gossip dialogues doesn't necessarily produce reliable facts. It means that for a large part social media analysis is built on quicksand.

Promoting impulsiveness and rudeness

Third, the instantaneous connections to other people that social media provide encourage people to send out their tweets, likes and comments to the world without hesitations. Communication via quick oneliners is fragmented, impulsive and superficial. It is likely to lack any nuance, consideration and depth. Addressing a virtual audience while lacking a clear, shared context may easily lead to unrestrained aggression, insults and threats. We know that public persons, such as politicians, CEOs or sportsmen are the easy targets of frustrated followers, who covered by anonymity bid against each other to express the greatest rudenesses. But we also know that the tirades and threats are

seldom turned into real actions. Social media talk may be spontaneous, and reveal hidden sentiments, but the excessive language should seldom be taken seriously.

Who's talking?

It is often claimed that social media allow us to get to know our customers better. Indeed, they allow us to closely follow customers, stakeholders, and target audiences and to learn about their vocabulary, profiles and preferences. However, this is not as easy as it looks like. Unfortunately we will only be able to collect partial data and compose partial profiles. Despite all the data we have available, identifying “the” customer is impossible. Even worse, the characteristics of targeted individuals are becoming intangible, volatile, fluid and unpredictable. Exactly the anonymity of the internet and the possibility of adopting different user profiles challenge individuals to explore different roles and switch frequently between those. More and more life is becoming a role play, a theater, allowing people to express and perform multiple personae. People are encouraged to play with their profiles, change and adapt these and pretend to be someone else without being essentially connected to their peer groups, thus ridiculing the very idea of target groups and the associated categorisations, subcultures, habits and preferences that marketers aim to identify. The goldmine may thus easily turn into a pitfall. Just as market researchers will be fooled by volatile individuals, the individuals will fool themselves. In addition, the emergence of artificial agents that cannot be distinguished from biological agents will further blur the playing field.

Nonsense and unruly behaviours

Now that social media users have learned that Big Brother NSA is watching us and every action and conversation is being tracked, traced, recorded and checked, they will return like for like by playing the fool. It's so easy to create a fake identity: pretending to be a brain surgeon, a millionaire or an artist rather than something less impressive. All security signals will change to red when we would collectively talk about “bombs” or “attacks”, if only for fun. The Uncyclopedia initiative, which started in 2005, deliberately publishes nonsense, fake articles, bad jokes and parodies. Its goal is to make available non-information, unjust statements, flagrant lies, deceit and politically incorrect information to everyone. Likewise, Deletionpedia specialised in recovering pages that were deleted from Wikipedia for reasons of “being not notable” or “manipulation by political and business interests”. Stealth data collection is a powerful incentive for promoting nonsense.

Abandoning the instrumental view

It is important not to view social media as a simple measurement instrument to be used for market research. Social media transcend the level of mere data collection in that the “instrument” of data collection coincides with the object of investigation, which is an organism of networked people. Hence, social media are both the camera and the scene to be captured; they are both the microphone and the sounds to be recorded. In addition, most social media users are well aware (or should be aware) of the fact that their data are being tracked and used for distilling profiles and preferences and thus may expressly play a trick on marketers by responding or behaving in unexpected ways (a game-theoretical mechanism: we know that they know that we know that ...etc). Since most social media platforms offer open interfaces or built-in analytics packages, users and user groups will be able to collect and evaluate the social media data themselves, and adjust their behaviours in preferential ways. This causes a self-establishing feedback loop, which is likely to amplify self-fulfilling prophecies and fallacies rather than real and trustworthy phenomena. Marketers could benefit from the same tricks (and they do!) by engaging in social media and by

influencing markets, while they adjust the very nature of the networks that they are part of. Still, it is hard to become a successful crystal-gazer and foretell the future. In 1999 no one anticipated the social web and the huge impact it would have on society and culture. I'm not sure if social media will in the end make a difference in foretelling the future. I suppose it still can be a moneymaker.