

# WHAT OPINIONS ARE FLOATING AROUND THE INTERNET? ANALYZING CUSTOMER SENTIMENT

Sentiment analysis and opinion mining seems to be a hot topic these days. Together with the spontaneous and massive expansion of new media (esp. social networks) this issue is being perceived as one of the major trends for the nearest future. Google Trends (a tool showing how often a particular search-term has been entered across various regions of the world and in various languages) reports a fivefold increase of search queries since 2007. However opinion mining and sentiment analysis are not only a recent phenomenon related to the development of information technology. A long time before the World Wide Web was born, our attitudes or (consumer's) decisions had been influenced by our relatives or friends around us. As we specialize and narrow our focus, the more frequently we need additional information about certain fields. Thus we have been looking for specialists or professionals with (life) experience who are able to share their opinion or advice. A car enthusiast whom we know could recommend a mechanic or a garage, a respectable person we appreciate could help us with the decision for whom to vote in the local election, or a colleague might provide a reference about a job applicant we would like to hire for our company.

Before the mass media has evolved, people were able to receive an input from just limited surrounding. This neighborhood was more likely limited to the geographical area where the individual lived and to the community they belonged to. As the TV and radio became widespread, the range of people who could influence public meaning greatly expanded. The distance is no longer a barrier and audiences can be approached in many ways. Persons speaking to the audience through the broadcast became strong opinion makers influencing and changing public attitudes and beliefs. Besides the *crème de la crème* of world politics (i.e. Barack Obama, Mario Monti, Mitt Romney or Hillary Clinton) and CEOs of the largest companies (i.e. Tim Cook from Apple, Sheryl Sandberg from Facebook or Virginia Rometty from IBM) the Time Magazine's List of The World's 100 Most Influential People from 2012 also includes sport players (Jeremy Lin, Novak Djokovic, or Lionel Messi) artists (Rihanna, Adele) or even famous Chefs (Jose Andres). Many of those "celebrities" do influence masses in the areas far beyond their expertise.

The online environment made it further possible to find out what others are thinking or experiencing, no matter if those are our personal contacts or well known professionals. In 2008, Pew Internet & American Life Project Report discovered that more than 80 percent of US internet users have previously done some online research about a product. Twenty percent do so commonly. More than 75 % of online-hooked customers confirm that reviews have a significant influence on their purchase and they are willing to pay more for a product

with better customer reviews. In addition, one third of users has posted an online review or rating regarding a product or service and thus became an influencer themselves.

Therefore continuous and systematic new media monitoring became an integral part of company processes. The aim of such activities is to recognize whether and (if so) in which context the media speaks about the company, hence companies can adjust their strategies and react in advance according to public opinions or attitudes. George Washington, the first President of the US, at the end of 18th century, monitored newspapers to monitor attitudes of his fellow citizens and what they were saying about him. Many of his successors also analyzed their media picture and 19th century witnessed an explosion of new companies scanning the media and cutting out news, providing the reports to government institutions, companies or individuals.

First press cutting (called "clipping" in the US) services retrieved information from media manually. They employed human readers to scan newspapers and look for keywords of (multiple) clients. Even if such services served only to a limited geographic area (mostly just one country) and monitored the media in just one single language, the human readers typically missed one third of articles containing keywords (as a result of not reading thoroughly but only fast scanning the articles). In the 1950s companies started to monitor TV broadcasts. Again, contracted personnel watched all programs and monitored broadcasted contents for keywords. With the introduction of Betamax and VHS technology, the monitoring process enabled recording and post-processing. Moreover, first automated speech-to-text systems have been developed, although they were still only 70–80 % complete and accurate.

The introduction of the World Wide Web in 1990s brought revolutionary changes on the media market. Almost all media are now digitalized and available online. New monitoring companies offer a computer-based processing where country borders or national languages do not matter anymore. CyberAlert – the world press clipping service – monitors 55 000+ online news sources in 250+ national languages 24 hours a day. New websites spontaneously emerge, change and disappear every day. A common user is not only a web content consumer, but also contributes as an author. We write product reviews, post comments in discussion forums or on social networks or even have our personal online blogs. Such fragmented and unofficial information is a significant source of word-of-mouth or "buzz" about companies and products. Many of today's most influential blogs (i.e. The Huffington Post, Techcrunch, Engadget, Mashable) began just few years ago as a one-man-show, quickly growing into internationally recognized and respected media.

Media monitoring has rapidly changed. Recognizing sentiment and finding out about people's attitudes still remains a challenge. However, utilizing sentiment analysis, companies can gain advantage from the insight in (social) media, recognize company or product reputation or develop a new marketing strategy addressing the negative sentiment and positively impacting on consumer's perception. Moreover, top influencers and opinion makers can be identified. Sentiment analysis may be also useful in other domains for predicting the result of (presidential) election or pointing a virtual finger at the next American Idol.

Old-fashioned press cutting services just monitored the media and did not supply any interpretation of the outputs provided. Even these days when the data are collected automatically, companies mostly interpret them manually. Such a random approach is then susceptible to subjective error and is dependent on the experience and skills of the persons

performing the analysis. Nevertheless, computer-based processing and modeling allows for automated sentiment analysis. Automated systems are repeatedly able to determine the sentiment (whether positive or negative) with 70–80 % accuracy. If automated systems were absolutely correct about sentiment classification, humans would still disagree with the results about 20 % of the time which is in no contrast to the level of agreement between raters they reached in human-rated analyses.

Automated sentiment analysis became an integral part of market research activities in companies such as Kia Motors, Best Buy, Deutsche Bank, Southwest Airlines or Paramount Pictures. Although the analysis cannot provide a hundred percent certainty, it can definitely offer important insights. It could be used to gain understanding of customer experience, recognize competitive dangers or discover emerging market opportunities. Opinions floating around the cyberspace begin to represent vox populi to the degree, in which most market segments participate in online discussions. Automated sentiment analysis might be a good way how to deal with cluttered online environment and how to get insight into the mind of a new millennial consumer.

#### Résumé — AKÉ NÁZORY PRÚDIA OKOLO INTERNETU? ANALÝZA SENTIMENTU ZÁKAZNÍKOV

*Analýza sentimentu a hĺbková analýza názorov sú, zdá sa, horúcou témou dnešných dní. Spolu so spontánnym a masovým rozšírením nových médií (najmä sociálnych sietí) je táto otázka vnímaná ako jeden z hlavných trendov blízkej budúcnosti. Google Trends (nástroj, ktorý ukazuje ako často bol určitý termín na vyhľadávanie v rôznych regiónoch a rôznych jazykoch zadaný) hlási od roku 2007 pätnásobný nárast vyhľadávacích dopytov. Avšak hĺbková analýza názorov a analýza sentimentu nie sú len novým fenoménom, ktorý súvisí s rozvojom informačných technológií. Dávno pred narodením World Wide Web ovplyvňovali naše postoje a (spotrebiteľské) rozhodnutia naši príbuzní a priatelia v našom okolí. Čím viac sa špecializujeme a zužujeme naše zameranie, tým častejšie potrebujeme dodatočné informácie o určitých oblastiach. Preto vyhľadávame špecialistov a profesionálov so (životnou) skúsenosťou, ktorí sú schopní s nami zdieľať svoje názory a rady. Automatizovaná analýza sentimentu sa stala integrálnou súčasťou aktivít marketingového výskumu v spoločnostiach Kia Motors, Best Buy, Deutsche Bank, Southwest Airlines alebo Paramount Pictures. Aj keď analýza nedokáže poskytnúť stopercentnú istotu, určite dokáže ponúknuť dôležité informácie. Pomáha nám pochopiť skúsenosť zákazníka, uvedomiť si konkurenčné nebezpečenstvá a objavovať vynárajúce sa trhové príležitosti. Názory prúdiace okolo kybernetického priestoru začínajú reprezentovať hlas ľudu v takej miere, v ktorej sa väčšina trhových segmentov zúčastňuje online diskusií. Automatizovaná analýza sentimentu môže byť dobrým spôsobom ako sa vysporiadať s preplneným online prostredím a ako získať pohľad do toho, čo sa deje v hlave spotrebiteľa nového tisícročia.*

**Kontakt na autora | Address** — Ing. Tomáš Kincl, Ph.D., Škoda Auto a. s. Vysoká škola, Katedra managementu a marketingu, Tř. V. Klementa 869, 293 60 Mladá Boleslav, Česká republika, e-mail: ykincl@is.savs.cz