Emotions in Citizens' Comments on the Internet as Predictors of Election Success

Danira Matijaca[*]

SUMMARY

Citizens express attitudes toward politicians and policies as special kinds of judgments colored by emotions and we call them sentiments. In this way, they participate in the process of creating public opinion, which can be manifested in different ways. Nowadays, it could be expressed through citizen comments on Websites, such as news portals, blogs and forums. Sentiment analysis of comments can give us an insight into the fundamental beliefs of citizens and their political profiles, which is very important for politicians when they want to explore public opinion, especially at the time of elections. This method is based on the extraction of sentiments expressed by words and terms in citizen comments on the Web. This paper presents the results of automated or manual sentiment analysis of citizens' comments about politics on news websites in different countries. It shows that citizens generally have a dominantly negative public opinion on politicians and politics and that different news websites gather specific political profiles of citizens. The comparison between the characteristics of public opinion and election results suggests that the results of sentiment analysis can serve as predictors of election success.

---

[*] Danira Matijaca, PhD student in doctoral study „ Society of Knowledge and Information Transfer" Department of Information Sciences, University of Zadar,
e-mail: danira.matijaca@xnet.hr

Introduction

Public opinion is a matter of huge interest to politicians, journalists and citizens themselves. For that reason, various methods are used to study its content and impact on politics. These complex studies include the analysis of different types of content in which public opinion is expressed and the correlation between public opinion and political actions. Public opinions express citizens' sentiments, defined as "attitudes colored by emotions" (Boiy et al., 2007: 350), and in regard to politics these sentiments refer to politicians, political decisions and events. Thus, one of the methods used for studying public opinion on politics is the analysis of sentiments that citizens express toward politicians and certain politics. Since the Internet appeared, the availability and amount of such contents, especially texts, have rapidly increased, because a large number of news websites, forums and news portals, allow citizens to leave comments.

Sentiments expressed in the text can be displayed by positive or negative numbers which represent the semantic orientation of text SO (Xu et al., 2012). Positive values express positive sentiments and negative values express negative sentiments, while the numerical value itself reflects the strength of sentiments. Computer programs can carry out analysis automatically, which is particularly useful when a large amount of text needs to be analyzed quickly. However, it is less precise than a manual analysis in which the researcher evaluates the strength and the sign of the sentiment expressed by a word or term (Caliandro & Gandini, 2016). Manual sentiment analysis represents a gold standard for calculating SO. The accuracy of computer programs for automated sentiment analysis is tested by comparing the results of computer analysis to the results obtained by manual analysis, since the computer program is not as good as man in perceiving irony and considering the context of the text (Tumitan & Becker, 2013). This paper explains the terms "public opinion", "sentiment analysis" and "semantic orientation", and shows some results of sentiment analysis of citizens' comments on the Internet. It also questions whether citizens' comments on news websites can be good indicators of public opinion for the whole population and, therefore, good predictors of election results.

Definitions and types of public opinion

The term "doxa" was used in Ancient Greece to describe what we call today "public opinion" and it referred to an uncertain judgment whose validity should be verified. Aristotle (Meta 1039b, as cited in Peters, 1967) defines doxa as "that which could be otherwise" and Plato (Republic V 474b – 480a, as cited in Turnbull, 1978: 9) describes it as an opinion which could be false or true, some „sort of propensity or

settled disposition". The term "opinion" was also used throughout history to determine one's reputation - the image that a person represents in the eyes of others (Habermas, 1991: 69).

The term "public opinion" as we understand it today appears in the 18th century, when all informed citizens started creating it. Being informed meant having enough free time to contemplate about politics so the citizens of the lowest strata of society who performed heavy physical work were not involved in the process. (Habermas, 1991: 102). Habermas (1996: 360) describes public opinion as the result of argumentative conversations in the public sphere, and the role of the political public sphere itself is "to ensure the formation of plurality of considered public opinions ". (Habermas, 2006: 416).

For Habermas (2006: 417), journalists, politicians, special interest groups and civil society members produce dominant public opinion. Insisting on argumentative conversations that precede the formation of public opinion, Habermas highlights the possibility of the existence of high quality and low quality public opinion, where quality is measured by the degree of the argumentation in conversations in the public sphere.

Tanta (2007: 15) does not define public opinion, but rather differentiates the terms "opinion" and "public". He defines "opinion" as "an insufficiently reasoned judgement" that expresses views on a subject that, when strong enough, leads to verbal and physical actions". He defines "public" as "a group of people who share common interest in a subject ". While Tanta highlights insufficient reasoning as the key feature of opinion, which is closely related to the understanding of the concept in Ancient Greece, some authors do not refer to the questions of grounding, argumentation or reasoning in the definition of public opinion.

For example, Key (1961, quoted in Zaller, 2003: 2) defines public opinion as "opinions held by private persons which governments find it prudent to heed ", thus giving it the key role in the process of making political decisions.

The political relevance of public opinion is also recognized by Erikson and Tedin (2015: 8) who say that public opinion reflects the wishes of adult citizens regarding the issues that are important to the government, and they describe public opinion as the verbal expression of the attitudes. When these attitudes are colored by emotions we can talk about sentiments, since Boiy et al. (2007: 350) defined them as judgments or ideas influenced or colored by emotions. Because they are related to emotions, sentiments are necessarily subjective.

Liu (2012: 19) believes that sentiments are constituent parts of an opinion. He identifies five elements of opinion: entity, aspect of the entity, sentiment, sentiment carrier, and time when sentiment is expressed.

An entity is an object toward which the sentiment is expressed and it can be a product, a service, a person, a topic, or an event. The aspect of the entity is a feature or a category of entity for which citizens express sentiments. In his description, Liu (2012) sees emotion as the narrowest concept included in the wider concept of sentiment, and opinion as an even wider concept that includes the sentiment, the person who expresses it, the entity, the aspect of entity, and the time when sentiment is expressed.

Entman and Herbst (2001: 206-209) make a distinction between active public opinion, mass public opinion, latent public opinion and perceived public opinion. A smaller number of citizens who actively participate in civil society and politics, have firm attitudes and are not subjected to manipulation, create active public opinion. Mass public opinion is the sum of the attitudes of all the citizens, many of whom are poorly informed, influenced by media, governments, events and surveys, so it can be said that such public opinion is subjected to manipulation. Perceived public opinion is created and empowered by media and it refers to the assumed attitudes of citizens. For this type of public opinion, Kuhnen and Niessen (2010) used the term "widespread opinion". There is another type of public opinion that is of particular interest to politicians, and Key (1961: 262, quoted in Zaller, 2003: 2) called it latent public opinion. It refers to the future attitudes of citizens and it manifests itself at the time of elections as a response to political actions.

The aforementioned definitions presume the involvement of all constituents of society in the process of creating public opinion that is composed of citizens' attitudes and feelings. However, there is another concept of defining public opinion, according to which public opinion constitutes attitudes stemming from reasoned conversations of parliamentary representatives (Habermas, 1991: 238).

Role of public opinion in the modern society

In any case, whether expressed by citizens or parliamentarians, in modern society public opinion becomes a powerful participant of the political public sphere (Krippendordf, 2005: 129) and can serve as a mechanism of governing (Kuhnen & Niessen, 2010: 29). Guizot (1852, quoted in Habermas, 1991: 101) writes about the rule of public opinion, typical for social systems that encourage citizens to constantly seek truth and justice. In such societies, what is true and just regulates the behavior of government, and according to Guizot, this can be achieved in parliamentary democracies. Namely, if conversations in parliaments fit the Habermas's (2006) description of deliberation, which means they are a reasoned with valid arguments, then they can lead to consensus that is subsequently transformed to political decisions. Furthermore, deliberations in parliaments are public because citizens can fol-

low their course through the media. In addition, journalists themselves encourage citizens to participate in public deliberations and present their views to politicians, in order to contribute to the search for what is true and best for all. Nowadays, the Internet platform offers opportunities for realizing the rule of public opinion. There are Web sites, such as news portals and forums, where citizens can freely share their opinions and engage in conversations with each other, and their conversations are visible to the public. Public opinion is also manifested at election time through the voting system and it can be expressed numerically by numbers of the mandates won by each party (Gerhard, 1993: 26, in Habermas, 2006: 418). However, one must perceive the two-way influence of public opinion and politics. Namely, politicians can also influence citizens by taking positive, negative, or neutral attitudes toward their opinions, and encourage them to reconsider their views. In this way, communication between the center of the political sphere, in which politicians are, and its periphery, in which citizens are, takes place to filter high-quality public opinion only, the one which is created in the process of public deliberation. This kind of communication creates a framework for making legitimate political decisions (Habermas, 2006: 417-418).

The mutual influence of politics and public opinion has been confirmed in many researches by measuring the responsiveness of politics to public opinion. It was found that there is a correlation between changes in political decisions and changes in public opinion (Manza & Cook, 2002: 632), but the influence of public opinion on politics is greater than the influence of politics on citizens' attitudes (Page & Shapiro, 1983: 186-187). However, in order to achieve the rule of public opinion, two Habermas's (2006: 420) criteria have to be met: an independent media system that offers a public deliberation platform, and a strong civil society that encourages citizens to participate in public deliberations. Considering the definitions of public opinion so far described, we can say that a number of its features are hardly realized in practice and therefore we should rather talk about an ideal that needs to be sought after. First, we are faced with the questionable independence of the media that can be influenced by power structures and this drawback hinders the process of qualitative and freely shaped citizens' attitudes in public deliberations (Habermas, 2006: 418). Namely, political elites become less sensitive to public opinion, thus more powerful, whenever citizens are poorly informed, manipulated by media, and rarely participate in public deliberations (Fishkin et al., 2000: 659). Furthermore, politicians and different interest groups can lie and deceive citizens, and after that make decisions based on manipulated public opinion, which is contrary to democratic principles (Page and Shapiro, 1983: 189).

Erikson and Tadin (2015) distinguish public opinion from its indicators: polls, demonstrations and similar gatherings of citizens, and such a differentiation is consid-

ered necessary due to the possibility of misinterpretation of the results and because the indicators cannot reveal the attitudes of the entire population. The results of polls, as the most common indicators of public opinion after World War II, represent a set of individual responses of citizens and express mostly mass public opinion (Entman & Herbst, 2001), while others' indicators are rarely taken into account: public gatherings, strikes, social movements and letters of citizens sent to newspaper editors and political elites (Manza & Cook, 2002: 632). Indicators of perceived public opinion can be found by content analysis of newspaper content, and in the case of latent public opinion we cannot talk about indicators, rather about predictors that politicians note by talking to colleagues, thanks to the experience in politics and by instinct – sense for the citizens' preferences (Entman & Herbst, 2001: 207-208).

Analyzing public opinion on the Web

The emergence of the Internet has enabled the creation of a new platform for the development of electronic democracy in which citizens express their opinions through new technologies. We use the term "e-citizens" for citizens active in the internet's environment (Clift, 2003) and their comments on the news websites represent a rich source for revealing public opinion (Liu, 2012: 5). The citizens' interest in comment writing can lead us to the conclusion that such comments express an active public opinion which is not subject to manipulation. However, that conclusion is valid only on the assumption that citizens are well informed about what they are writing. In addition, such texts can express citizens' fundamental beliefs and their set of values, enabling us to anticipate latent public opinion.

When emotions, as constituent parts of sentiments contained in an opinion, are expressed through a text, words become a means of diagnosing the mental, social and physical conditions of the person who says them (Pennebaker et al., 2003: 548). This is the reason why a novel scientific technique- sentiment analysis or opinion mining is being developed and it is conducted by a computer program (Jakopović & Preradović, 2016: 68). Sentiment analysis can also be performed manually. In that case, a researcher finds and denotes a piece of text, which expresses a sentiment, as positive, negative, or neutral. Manual analysis is more precise, especially, when the sentiment is expressed with irony and when it is important to determine the context of the written text. Therefore, it is the golden standard for sentiment analysis which is used for measuring the precision of automatic analysis (Hatzivassiloglou & McKeown, 1997: 175). However, manual analysis requires a lot more time so, it is appropriate to use automatic analysis when we have to handle a large amount of texts archived on the news websites (Caliandro & Gandini, 2016: 198). It is necessary to keep in mind that a text contains physical (objective) and semantic (subjec-

tive) information. Physical information refers to what we really see; elements such as a word in a sentence or the word order in a sentence and because semantics investigates „what we communicate to our species using the language as a tool of communication" (Diamant, 2012: 1) semantic information is obtained by interpreting physical information (Diamant, 2012: 7). Semantics investigates what meanings refer to, how they are associated with words, phrases, and sentences, and how we can use and combine meanings for concluding and thinking (Beaver and Freeze, 2011: 1). Therefore, getting the most accurate semantic information is crucial to sentiment analysis. Yehoshua Bar-Hillel and Rudolf Carnap (1952, quoted in Diamant, 2012:2) were the first to use the term "semantic information" and they defined it as the meaning we give to the objective information. Diamant (2012: 5) argues, rightly, that semantic information is in the observer's head and it is, therefore, obvious why manual analysis is the golden standard. Indeed, computer software is only more or less successful in trying to emulate a person.

A sentiment in a word or in a phrase, as an emotional attitude, is defined by two pieces of semantic information: the strength of the sentiment and the sign or the polarity of the sentiment, which can be positive or negative. This information can be expressed with a positive or with a negative number, which, in that case, represents the semantic orientation of a word or a phrase and is denoted by SO. A negative sign is attached to a word or a phrase that expresses a negative sentiment, and a positive sign is attached to a word or a phrase that expresses a positive sentiment (Xu et al., 2012: 279). The semantic orientation of the whole text is obtained by summing up the SO of the words or phrases in the text that express sentiments. Computer programs for sentiment analysis "translate" words and phrases into positive or negative numerical values and that process is based on the physical information, such as the appearance of a certain word, the type of linkers between words and the word order, in the text.

Ding et al. (2008) created a computer algorithm that, by means of a pre-made lexicon, recognized sentiment bearers in a text – opinion words and phrases that express sentiments as well as opinion idioms. The sentiment bearers were assigned the value of +1 (for a positive sentiment) or the value of -1 (for a negative sentiment), and the total SO of the analyzed text was obtained by summing up the SO of all sentiment bearers in a text.

Determining the SO of the citizen comments on the news websites can serve as a method for investigating the public opinion on politicians and certain policies. Park et al. (2011) noted that citizen comments could be positive, negative or unclear. Positive comments are expressed through praise and optimism while negative comments are expressed through ridicule and criticism. If we want to carry out manual or automatic analysis, it is possible to apply a simple basic formula for determining

the SO of the citizen comments that refers to certain politicians or policies (adapted from Ding et al., 2008):

$$SO \ (politician \ / \ policy) = \sum_{i}^{n} SO \ (w_i)$$

The SO (politician / policy) is the total SO of a citizen's comment; n represents the total number of sentiment bearers in the comment; and SO (w_i) represents the semantic orientation of the individual sentiment bearer in the comment.

Finally, we can say that public opinion about a politician or a policy can be carried out by the sentiment analysis of citizen comments on the news websites (Abdulla et al., 2014; Park & Cardie, 2014; Ruiz et al., 2011; Noci et al., 2012, 2013, Tumitan & Becker, 2013, Shumueli et al., 2012; Meguebli et al., 2014). Analyzing sentiments means getting the semantic information from the comments and expressing it with numeric values. At the core of such an analysis, conducted by a researcher or by a computer program, is the assignment of positive or negative numbers to words and phrases that express sentiments and the calculation based on these numbers to determine whether the comments about a politician or a policy are positive, negative or neutral.

Previous sentiment analysis of citizen online comments about politics

Sentiment analysis of citizen comments on different news websites can also be used to determine a political profile of the citizens. This was confirmed by Meguebli et al. (2014) who designed a methodology for determining the political orientation of the citizens based on the sentiment analysis of the comments on the news portals. Admittedly, determining the political orientation of the citizens was only a means of evaluating the accuracy of the computer algorithm for automatic sentiment analysis. Nevertheless, they offered a methodology for that kind of profiling. They analyzed the comments of 500 citizens; 290 of them were the visitors of the news portal CNN and 210 of them were the visitors of the Egyptian news portal Al Jazeera. The key to this research was to analyze comments in the four-year period; from October 2009 to September 2013. In the first step, they categorized political topics the citizens spoke about and then they determined how many citizens expressed positive, negative and neutral sentiments on each topic. The political profile of a citizen was described based on a series of topics and the positive, the negative or the neutral sentiments signs associated to them. Five students conducted the analysis and the units of the analysis were all the sentences of a single citizen collected during the research period. The words and phrases that expressed sentiments were detected in

each of the sentences by means of the previously created lexicon. These sentences were designated as positive if the sum of the SO of the words and phrases that express sentiments was positive while if the sum was negative, the sentences were designated as negative. The prevailing citizen sentiment toward a topic was determined by summing up the SO of all the sentences written on that topic. For example, a typical topic that allows detecting a citizen's political profile is the right to abortion. Liberals express positive sentiments and conservatives express negative sentiments toward this topic. A list of topics and the associated prevailing sentiments was created for each citizen and, based on that, the citizen was labelled as a Democrat or a Republican(if it was a CNN portal visitor) and as a secularist or an Islamist(if it was a visitor of the Egyptian portal).

Sentiment analysis of citizen comments on the news portals is particularly useful at the time of a pre-election campaign when it can predict the outcome of the elections, as there is a correlation between the results of the sentiment analysis and the results of the pre-election polls.

Table 1   Ratios of positive and negative sentiments in the comments
            on the Brazilian portal
Tablica 1 Omjer pozitivnih i negativnih emocija u komentarima na brazilskom portalu
*   Folha de São Paulo (Tumitan and Becker, 2013:4)*

Description	Formula		Haddad	Serra	Russomanno
Ratio of positive sentiment of an entity to the negative sentiment of the same entity	$s_d = \dfrac{pos_e}{neg_e}$	(1)	0.57/0.44	0.54/-0.14	0.12/-0.04
Ratio of positive sentiment to the total sentiment	$s_d = \dfrac{pos_e}{pos_e + neg_e}$	(2)	0.56/0.40	0.54/-0.16	0.08/-0.04
Ratio of negative sentiment to the total sentiment	$s_d = \dfrac{neg_e}{pos_e + neg_e}$	(3)	-0.56/-0.40	-0.54/0.16	-0.08/0.04
Ratio of positive sentiment of an entity to the positive sentiment of all entities	$s_d = \dfrac{pos_e}{pos_{entities}}$	(4)	0.09/0.07	0.22/0.29	0.39/-0.29
Ratio of negative sentiment of an entity to the negative sentiment of all entities	$s_d = \dfrac{neg_e}{neg_{entities}}$	(5)	-0.35/-0.23	0.16/0.34	0.16/-0.04

Folha de São Paulo (Tumitan and Becker, 2013:4)

Table 1 shows the results of the research conducted by Tumitan and Becker (2013). They analyzed sentiments in the comments on the Brazilian portal Folha de São Paulo, written during the pre-election campaign, before the first election round for the mayor of São Paulo. Their research included 36,108 comments on 583 news articles written about the three leading candidates and they extracted 9,758 sentences in which these candidates were mentioned. To evaluate the accuracy of automatic analysis, they first conducted the manual analysis of 600 sentences that were selected by random sampling. Then, they created several computer algorithms to determine the SO of the sentences and selected the most accurate one, based on the comparison of the results of the automatic analysis with the manual one. After they have finished the automatic analysis of all sentences in the sample, they created a table for each candidate with the results of the analysis on a daily basis. The tables showed:

a) the ratio of positive and negative sentences
b) the ratio of positive sentences and the total number of emotional sentences
c) the ratio of negative sentences and the total number of emotional sentences
d) the ratio of positive sentences and the total number of positive sentences about all candidates
e) the ratio of negative sentences and the total number of negative sentences about all candidates

These ratios were compared with the results of the pre-election polls, conducted during the research period and, by using Pearson correlation coefficient, it was found that there was a high to moderate correlation between the ratio of positive and negative sentences and the results of the pre-election polls. This means that the increase in the number of positive sentences about candidates most frequently followed the increase in the number of voters inclined toward them. During the manual analysis, 80% of sentences were negative, 12% were positive and 7% were neutral. Similar results were obtained after the automatic analysis. Based on these results Tumitan and Becker (2013, p.6) concluded that most comments express negative sentiments and frustration about politics.

Carvalho et al. (2011) analyzed the citizen comments on the Portuguese news portal Público, written about ten news articles about TV debates ahead of the election for the Portuguese Parliament in 2009. Their results lead to the conclusion that the key is not so much the polarity of the expressed sentiments, but the existence of emotions in the comments and that the most successful candidate in the election was the one about whom the largest number of negative sentences was written. They determined the SO of the comments related to the candidates of the largest Portuguese parties and their sample encompassed 2,795 comments and approximately 8,000 sentences. Given the expressed sentiments, each sentence could get a numerical

value: -2, -1, 0, 1 and 2, where positive values were assigned to the sentences with positive sentiments and negative to those with negative sentiments. The results showed that 60% of the sentences expressed a negative opinion, 15% expressed a positive opinion, 13% expressed a neutral opinion and 14% of the sentences expressed no opinion or they were unclear. For one of the politicians, there was a similar number of positive and negative comments while negative comments were prevalent for the rest of them. It is interesting that the largest number of negative comments was written about a candidate who was chosen for the prime minister twice, including in 2009, and the smallest number of negative comments was written about the candidate who received the lowest percentage of votes in the 2009 elections. There was a high correlation between the number of votes received by a single candidate and the total number of written comments about that candidate. Taking these results into account, we can conclude that the sign of the sentiments expressed in the comments is not crucial for the success of candidates and that the number of comments about them is much more important. The citizen comments in this survey were mostly negative. Therefore, Carvalho et al. (2011: 568) pointed out that citizens prefer to react when they disagree with politicians rather than when they agree with them. Expressing political frustration on the news websites, such as newspaper portals, is further facilitated by the fact that citizens may be anonymous while writing the comments. During the manual analysis, Carvalho et al. (2011) found 11% of ironic sentences that can easily become an error source in automatic analysis; due to the automatic recognition of words and phrases denoting positive sentiments such sentences are denoted as positive, although they are not. Therefore, according to Carvalho et al. (2011), in the future, most attention should be directed to the creation of computer algorithms for a more precise identification of positive sentences.

Obviously, semantic information in online citizen comments can not only be used to thoroughly investigate the public opinion about a politician or a policy but also to determine the political profile of citizens and to predict, more or less reliable, election results. The information about citizen sentiments in online comments also enables us to describe the characteristics of a communication specific to different news websites where citizens write their comments. Balasubramanyan et al. (2012) chose a period ahead of the 2008 presidential elections and automatically analyzed the citizen comments on five US blogs; some of which had a larger number of liberal blog users while others had a larger number of conservative blog users. A computer algorithm calculated the positivity or the negativity of the comments by detecting the words that express sentiments. For each blog, there was a pre-made lexicon with such words, typical of the community of citizens on a particular blog. The SO of the comments was calculated by summing up the SO of the automatically recognized words and the results of this research showed that there was a different ratio be-

tween the total number of positive and negative comments for each blog. One blog with a large number of liberal users and one blog with the majority of conservative users had more positive than negative comments. One blog with more users that are conservative and two with more users that are liberal had mostly negative comments. The dominance and the strength of positive or negative sentiments in the comments on a particular blog depends on the specificity of the users' community, whether it is predominantly conservative or liberal; the strength and the polarity of the sentiments also depend on the very subject of the citizens' comments. Furthermore, they noted that there was a connection between the strength of the expressed sentiment and the total number of comments on a subject, meaning that topics causing strong positive or negative sentiments also cause more comments.

Another example of manual sentiment analysis comes from Abdulla et al. (2014: 110). In their research, they included 6,921 comments and reviews written on the most popular Arabic social media - Yahoo! Maktoob and found a correlation between the strength and the sign of sentiments and the topics of the comments. The largest number of negative sentiments was expressed in the reviews and comments on politics (65, 5%) and the smallest number was expressed in the comments and reviews related to art (31, 43%). At the same time, politics also had the least positive comments (10, 36%) and art had the most positive comments (34, 4%). Political topics obviously do not leave citizens indifferent, since there was the smallest number of neutral comments about these issues (19, 35%). This research confirmed, once more, that political topics provoke citizens to express their attitudes more often when they are frustrated and when they disagree with something meaning such comments are filled with negative sentiments.

Conclusion

Public opinion as the main "product" of the public sphere is a subject of interest for the politicians, citizens and journalists in democratic social systems. It can be described as a wider concept that encompasses attitudes coloured with emotions, which Boyle et al. (2007) call sentiments. The ancient concept of opinion as a judgment that is uncertain is still present today as much as a citizen's opinion about politicians or certain policies is manipulated. The general lack of citizens' interest in engaging in civil society and citizens not being well-informed pave the way for manipulation. What Habermas (2006) represents as a desirable conversation between citizens and calls public deliberation presupposes good argumentation and a high level of being informed. That can ultimately lead to a consensus and the citizens' opinions, which have emerged through such a form of conversation in the public sphere, are most definitely to be considered high quality public opinions. We can define public opinion as a set of all citizens' attitudes about a topic and, if we

represent it as the one that defines the politics of those who rule, then we can say that it represents a prevailing opinion or the opinion of the majority. That is why mass public opinion, being subject to manipulation the most, has the most influence on politics; since in democratic systems all citizens can express their opinions about politics through elections every four years. Such opinion is revealed before elections by surveys that have been the most popular method for public opinion polling after the World War II. The problem with this kind of opinion is that a large number of citizens who shape it are not well informed and do not participate in something that Habermas (2006) defines as a public deliberation. It seems that such a desirable process of conversation can only be achieved with a small number of well informed and active citizens whose opinions are not subject to manipulation, and which, therefore, form the active public opinion. In addition to surveys, the content analysis of citizens' conversations could reveal the characteristics of active and of massive public opinion as well as the characteristics of latent public opinion, which is of particular interest to politicians. Namely, latent public opinion is the one to appear in the future, during the elections, and it reflects the underlying values and beliefs of the citizens. Politicians can predict such an opinion on the basis of instinct and experience and in addition to the content analysis of citizens' conversations, which are a rich source for the research of attitudes and fundamental beliefs of the voters; it is possible to analyze other forms of public opinion manifestation such as street turmoil, protests, social movements, etc. Since massive public opinion, which has the most influence on politics, is subject to manipulation by the mainstream media, we can notice one more kind of public opinion, the so-called perceived public opinion, which refers to the opinion empowered through media contents, and the analysis of such contents reveals its characteristics. Without discussing how well each public opinion is and how much it reflects the opinions of the majority or of only a small number of citizens, on the online news websites, such as news portals, blogs and forums, we can find a rich source of written texts for analyzing all kinds of public opinions. Content analysis of comments allows us to obtain semantic information about the sentiments that citizens express toward a particular politician or a policy. When semantic information in a text is displayed by positive and negative numerical values, we call it the semantic orientation SO of the text, and it expresses the sign and the strength of the sentiments expressed in the text. By calculating the SO of citizen comments on the news websites, we can determine the political profiles of the citizens, the characteristics of the individual news websites with regard to the characteristics of the citizens' conversations, and even highlight the predictors of the electoral results. The winners of elections are mostly those most comments, coloured with emotions, are written about even though the positivity or negativity of the emotions themselves is not so crucial. Moreover, it is possible that a candidate who receives the most negative comments ultimately wins and we can explain that taking into account research findings that showed the domination of negative emo-

tions in the citizen's comments about politics. Therefore, it is more than clear that citizens use Internet platforms mostly to express their disapproval and frustration about politics. A great number of citizens' comments on the internet represent a big challenge in the development of computer algorithms that could handle a large amount of texts in a short time during automatic sentiment analysis. The success of automatic sentiment analysis is measured by comparing its results with results obtained by manual analysis, which is conducted by a researcher and is considered to be the golden standard. In addition to improving automatic sentiment analysis, researchers are faced with the challenge of discovering a possible correlation between characteristics of all kinds of public opinions expressed in the comments on the news websites and the election results or the implementation of certain policies. Such an analysis can be used to measure a degree of policy responsiveness to the public opinion of those citizens who are active on the Internet and it could, thus, become a participant in the process of public opinion formation, as the process of shaping attitudes is exposed to the influence of all information, including those that concern public opinion itself. If we want to realize Guizot's idea about the rule of public opinion and if we want to stay committed to the goal of shaping quality opinions through public deliberations, thanks to automatic sentiment analysis, it is possible to approach Habermas' ideal of the public sphere and to provide quality opinion as the basis for making legitimate political decisions.

REFERENCES

Abdulla, N. A., Al-Ayyoub,M. & Al-Kabi, M. N., (2014), An extended analytical study of arabic sentiments, *International Journal of Big Data Intelligence 1, [Online], 1 (1-2), pp.* 103-113, <raspoloživo na: https://www.researchgate.net/profile/Mohammed_Al-Kabi/publication/264832333_An_extended_analytical_study_of_Arabic_sentiments/links/54253a780cf26120b7ac7f46.pdf>, [pristupljeno 23.04.2015.].

Beaver, D. & Frazee, J., (2011), Semantics, The Handbook of Computational Linguistics, Mitkov, R. (Ed.), [Online], <raspoloživo na: http://semanticsarchive.net/Archive/DAzYmYzO/semantics_oup.pdf>, [pristupljeno 07.06.2014.].

Boiy, E. et al., (2007), Automatic Sentiment Analysis in On-line Text, *ELPUB,* pp. 349 – 360.

Caliandro, A. & Gandini, A., (2016), Qualitative research in digital environments: a research toolkit, Routledge.

Diamant, E., (2012), Let us first agree on what the term" semantics" means: An unorthodox approach to an age-old debate, arXiv preprint, [Online], <raspoloživo na: https://arxiv.org/pdf/1201.0328>, [pristupljeno 16.06.2014.].

Ding, X., Bing, L. & Yu, P. S., (2008), A holistic lexicon-based approach to opinion mining, *Proceedings of the 2008 International Conference on Web Search and Data Mining*, [Online], ACM, <raspoloživo na: http://www.cs.uic.edu/~liub/FBS/opinion-mining-final-WSDM.pdf>, [pristupljeno 05.09.2014.].

Entman, R. M. & Herbst, S., (2001), Reframing public opinion as we have known it, *Mediated politics: Communication in the future of democracy,* pp. 203-225.

Erikson, R. S. & Tedin, K. L., (2015), American public opinion: Its origins, content and impact, Routledge.

Fishkin, J. S. & Luskin, R. C., (2005), Experimenting with a democratic ideal: Deliberative polling and public opinion, *Acta Politica,* [Online], 40 (3), pp. 284-298, <raspoloživo na: https://www.uvm.edu/~dguber/POLS234/articles/fishkin.pdf >, [pristupljeno 11.04.2014.].

Habermas, J., (1991), The structural transformation of the public sphere: An inquiry into a category of bourgeois society, Cambridge: MIT Press.

Habermas, J., (1996), Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy, Communication Theory, [Online], 16, <raspoloživo na: http://blogs.unpad.ac.id/teddykw/files/2012/07/J%C3%BCrgen-Habermas--Between-Facts-and-Norms.pdf>, [pristupljeno 07.06.2014.].

Habermas, J., (2006), Political communication in media society: Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research, Communication Theory, [Online], 16 (4), <raspoloživo na: http://mt.educarchile.cl/mt/jjbrunner/archives/1-habermas_deliberation2006.pdf>, [pristupljeno 22.06.2014.].

Hatzivassiloglou, V.& McKeown, K. R., (1997), Predicting the semantic orientation of adjectives, Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics. Association for Computational Linguistics, *[Online]*, 40 (3), pp. 284-298, <raspoloživo na: http://www.anthology.aclweb.org/P/P97/P97-1023.pdf>, [pristupljeno 11.04.2014.].

Jakopović, H. & Mikelić Preradović, N., (2016), Identifikacija online imidža organizacija temeljem analize sentimenata korisnički generiranog sadržaja na hrvatskim portalima, Medijska istraživanja: znanstveno-stručni časopis za novinarstvo i medije, [Online], 22 (2), pp. 63-82, <raspoloživo na: https://hrcak.srce.hr/file/253147>, [pristupljeno 08.10.2017.].

Krippendorff, K., (2005), The social construction of public opinion, *Kommunikation über Kommunikation*. VS Verlag für Sozialwissenschaften, [Online], pp. 129-149, <raspoloživo na: http://repository.upenn.edu/cgi/viewcontent.cgi?article=1077&context=asc_papers>, [pristupljeno 18.09.2014.].

Kuhnen, C. M. & Niessen, A., (2012), Public opinion and executive compensation, Management Science, [Online], 58 (7), pp. 1249-1272, <raspoloživo na:

https://www.econstor.eu/bitstream/10419/69237/1/735594988.pdf>, [pristupljeno 21.05.2014.].

Liu, B., (2012), Sentiment analysis and opinion mining, *Synthesis Lectures on Human Language Technologies* 5 (1), pp. 1-167.

Manza, J. & Cook, F. L., (2002), A democratic polity? Three views of policy responsiveness to public opinion in the United States, *American Politics Research, [Online],* 30 (6), pp. 630-667, <raspoloživo na: http://www.cassr.as.nyu.edu/docs/IO/3858/A_Democratic_Polity.pdf>, [pristupljeno 12.11.2015.].

Page, B. I. & Shapiro, R. Y., (1983), Effects of public opinion on policy, *American political science review,* [Online], 77 (1), pp. 175-190, <raspoloživo na: http://www.uvm.edu/~dguber/POLS234/articles/page.pdf>, [pristupljeno 06.04.2014.].

Park, S. et al., (2011), The politics of comments: predicting political orientation of news stories with commenters' sentiment patterns, Proceedings of the ACM 2011 conference on Computer supported cooperative work, [Online], ACM, <raspoloživo na: https://pdfs.semanticscholar.org/c849/e41faef43945fcb02a-12bae0ca6da50fc795.pdf>, [pristupljeno 05.09.2014.].

Pennebaker, J. W., Mehl, M. R. & Niederhoffer, K. G., (2003), Psychological aspects of natural language use: Our words, our selves, *Annual review of psychology*, [Online], 54 (1), pp. 547-577, <raspoloživo na: http://psych.uw.edu.pl/mikk/4621%20METOD_COMM%20APS/dodatki/Our%20Word%20our%20selfes.pdf>, [pristupljeno 10.04.2015.].

Peters, Francis E, (1967), Greek philosophical terms: A historical lexicon, NYU Press.

Tanta, I., (2007), Oblikovanje mnijenja ili nužnost manipulacije, MEDIANALI-znanstveni časopis za medije, novinarstvo, masovno komuniciranje, odnose s javnostima i kulturu društva, [Online], 1 (2), pp. 13-32, <raspoloživo na: http://hrcak.srce.hr/file/65046>, [pristupljeno 08.10.2014.].

Tumitan, D. & Becker, K., (2013), Tracking Sentiment Evolution on User-Generated Content: A Case Study on the Brazilian Political Scene, *SBBD (Short Papers)*, [Online], <raspoloživo na: https://pdfs.semanticscholar.org/0d7b/5599aced755c02764711c49a1c43791fb654.pdf>, [pristupljeno 13.06.2015.].

Turnbull, Robert G., (1978), Episteme and Doxa: Some Reflections on Eleatic and Heraclitean Themes in Plato, The Society for Ancient Greek Philosophy, Newsletter 91., <raspoloživo na: https://orb.binghamton.edu/sagp/91>, [pristupljeno 23.01.2017.].

Xu, T., Peng, Q. & Cheng, Y., (2012), Identifying the semantic orientation of terms using S-HAL for sentiment analysis, *Knowledge-Based Systems* 35, pp. 279-289.

Zaller, J., (2003), Coming to grips with VO Key's concept of latent opinion, *Electoral Democracy, [Online],* pp. 311-36, <raspoloživo na: https://www.econstor.eu/bitstream/10419/69237/1/735594988.pdf>, [pristupljeno 21.05.2014.].

Emocije u komentarima građana na Internetu kao pretkazatelji izbornog uspjeha

Danira Matijaca

SAŽETAK

Građani izražavaju stavove prema političarima i politikama kao posebne vrste prosudbi obojanih emocijama koje nazivamo sentimentima. Na taj način sudjeluju u oblikovanju javnog mnijenja koje se može očitovati na više načina. Danas je moguće javno mnijenje izraziti kroz komentare građana na mrežnim mjestima poput novinskih portala, foruma i blogova. Analiza sentimenata u takvim komentarima pruža nam uvid u njihova temeljna uvjerenja i političke profile, što je osobito važno političarima kada u vrijeme prije izbora žele istražiti javno mnijenje. Takva se analiza zasniva na uočavanju sentimenata izraženih kroz riječi i izraze u komentarima građana na Internetu. U ovom su radu prikazani rezultati automatske i ručne analize sentimenata na novinskim portalima u različitim državama. Uočeno je da građani uglavnom imaju negativno javno mnijenje o političarima i politikama te da različiti novinski portali okupljaju građane različitih političkih profila. Na temelju usporedbi karakteristika javnog mnijenja i izbornih rezultata može se zaključiti da rezultati analize sentimenata mogu poslužiti kao pretkazatelji izbornog uspjeha.

Ključne riječi: semantička orijentacija, analiza sentimenata, javno mnijenje, izbori, komentari građana, novinski portali