DETECTING THE ONLINE IMAGE OF "AVERAGE" RESTAURANTS ON TRIPADVISOR

Hrvoje Jakopović

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

ABSTRACT *Collective intelligence can be interpreted as the actions of individuals that provide collective effects. In online spaces, the more user comments about a matter of discussion, the higher the potential that certain repeated points of view will be used as a story frame. This observation can be a very useful explanation for the value of user comments, reviews and the ratings in the field of public relations. Nowadays, it has become noticeable that many indecisive people who are thinking of buying a product or using a certain service rely on information left by users who already have some kind of experience with the product or service. This information has an effect on decision-making and taking action. In the case of contemporary PR, collective intelligence, facilitated through user comments/reviews, is involved in the image making process. This paper uses the idea of collective intelligence to measure restaurants' online image, using sentiment analysis to gain insight to users' attitudes and opinions. Image is interpreted as a short-term outcome of organizational activities that can be identified through individual attitudes and opinions in this study. The author uses sentiment analysis, the use of natural language processing applications, to examine user comments and reviews for restaurants in Dubrovnik rated as "average" on the website TripAdvisor. This paper tests the accuracy of sentiment analysis software, therefore the efficiency of automated sentiment analysis is compared to human sentiment analysis. The results indicate that sentiment analysis tools could be important instruments for the estimation of a positive, negative, or neutral sentiment and detection of organization's online image.*

KEY WORDS

COLLECTIVE INTELLIGENCE, ONLINE IMAGE, SENTIMENT ANALYSIS, USER COMMENTS, PUBLIC RELATIONS, TRIPADVISOR

Author Note

Hrvoje Jakopović :: University of Zagreb, Faculty of Political Science, Croatia **::** hrvoje.jakopovic@fpzg.hr

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

INTRODUCTION

In the information age, online users are in the spotlight as they carry great power. They can potentially influence numerous individuals and organizations with their online comments, reviews, ratings, and opinions. For that reason, public relations (PR) and marketing experts in their everyday practice analyze user-generated content (UGC). User comments/reviews are a significant part of organizational image as they mediate experience. Users like to share their good or bad experiences, but also hear about other, similar experiences from other users. The things they hear affect them and reflect on their comments. As users publish more information on a certain topic, they intensify the dominant story frame. The more users are involved in communicating about a certain topic, the sum of their reviews will produce a more insightful evaluation of a product, service, organization or person. The overall number of user reviews represented on a certain website captures a momentary online image. In that context a consistency issue exists because the reviews of the product, service, organization or a person are not equally positive, negative or neutral all the time. However, by publishing their comments and reviews the combined actions of users facilitate collective intelligence (CI). This diffusion of information, its multiplication of individual opinions and consequently the creation of an online image supports the concept of collective intelligence. James Surowiecki emphasizes that the smartest groups are those ones which consists "of people with diverse perspectives who are able to stay independent of each other" (2004: 41), even if some member of the group is irrational he will not make the whole group less intelligent.

This paper measures restaurants' online image based on the idea of collective intelligence. The author uses and tests sentiment analysis, which can be observed through the use of natural language processing applications, to gain insight into users' attitudes and opinions. Sentiment is based on polarities of positive and negative emotions (Thelwall *et al*., 2012). The author examines it with human sentiment analysis and compares it with computer sentiment analysis. This study validates the accuracy of sentiment analysis programs in relation to referent human coding. Data for analysis is extracted from the most popular travel site, TripAdvisor (www.tripadvisor.com), and consists of user comments/reviews for restaurants in Dubrovnik which are rated as "average." Considering the aims and scope of the article, collective intelligence is observed through user comments and through the image making process, therefore the online image, as one part of the product of collective intelligence, is measured through the sentiment expressed in user comments/reviews. This article includes a conceptualization of collective intelligence and online image. The author discuses the connection between these two phenomena, collective intelligence and image – often referred to as mental and attitudinal constructs made from the creative process of information selection, elaboration, embellishment and ordering (Lee *et al*., 2005, cited in Lee *et al*., 2014). The research tests the relationship between user reviews (sentiment/textual content) and user ratings (on a scale from 1 to 5). Furthermore, it examines the precision of sentiment analysis tools which offer new possibilities in online data analysis and measuring organization's online image.

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

INDUCING IMAGE FROM COLLECTIVE INTELLIGENCE

Ideas of Collective Intelligence

The term collective intelligence (Livingstone, 2016; Woolley *et al*., 2015; Peters and Reveley, 2015) is used to describe a phenomenon present in everyday life, but at first it is not so easily noticeable. As individuals, our acts are often determined through particular interest, and in many cases our individual efforts take part in achieving collective effects. For example, success of a football team and wining the match depends on the actions of every player in the team. The concept of collective intelligence is putting various beings in a system in which collaboration among them is happening, sometimes even unconsciously, and providing collective effects. The colony of ants described by James Surowiecki in his book *The Wisdom of Crowds* may be the perfect example of collective intelligence in natural life: "No one ant runs the colony. No one issues orders. Each individual ant knows, on its own, almost nothing. Yet the colony successfully finds food, gets all its work done, and reproduces itself." (2004: 40) Human communities are organized somehow differently, firstly, with the respect to individual intelligence and capability, which is why organizations employ and manage workers by hierarchies. With a glimpse on how human society works, we can use a simple interpretation for collective intelligence among humans, explained as a "groups of individuals doing things collectively that seem intelligent" (Malone, 2008: 1). Jean Pierre Lévy defines it as "a form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills" (1997: 13). In contemporary human society, collective intelligence represents activities of separated individuals who are providing collective effects. This concept is often associated and derived from the term "crowdsourcing" which implies the gathering of individuals in order to perform collective actions and achieve common goals. Tanja Aitamurto (2011) points out that in many crowdsourcing projects, the participants associate themselves with the tasks, and the result is an aggregation of the input of the participants. In this regards, collective intelligence is a potential consequence of crowdsourcing projects.

It is possible to observe the enhancement and evolution of collective intelligence through the rise of information and communication technology (ICT) and the development of cyberspace. From this point of view, the concept has a much wider implementation because Web 2.0 and ICT strongly affect everyday life in urban environments, transforming societies and human interaction. Furthermore, new ideas of collective intelligence are emerging – Wikipedia, search engines and social networking sites. Social networking sites are spaces where users interact and exchange information and knowledge. François Cooren's (2004) analysis of board meeting excerpts from a drug rehabilitation centre and his previous findings indicate that collective intelligence is mostly present in forms of conversational behavior. A great number of user generated content is publicly available on the sites, and it is constantly enriched by other UGC. The more users are active on a certain topic, the more a specific story frame is intensified. Users spread dominant frames as a product of collective intelligence and influence on other users. Charlotte N. Gunawardena *et al*., share a similar perspective and observe "social networking technology as tools that facilitate collective intelligence through social negotiation when participants are engaged in a common goal or a shared practice"

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

(2009: 6). Henry Jenkins observes collective intelligence in terms of media consumption and recognizes how online fan communities are a type of "self-organizing group focused around the collective production, debate, and circulation of meanings, interpretations, and fantasies in response to various artifacts of contemporary popular culture" (2002). On the other hand, users are not always aware of their participation in facilitating CI. Marco Leimeister (2010: 245) points out that if we etymologically decompose collective intelligence, it is possible to deem that the term 'collective' describes a group of individuals who are not required to have the same attitudes or viewpoints. This thought is strengthened by Satnam Alag who sees collective intelligence as "powering a new breed of applications that invite users to interact, contribute content, connect with other users, and personalize the site experience" (2009: 18). One of the specific aspects of Web 2.0 based collective intelligence is also the disappearance of spatio-temporal boundaries, which in the past represented greater barriers than they do today. Cano Viktorsson (2013) shows how a Swedish radio station organized the exchange of traffic information before the massive use of the Internet as a way of collective intelligence in action, and he deemed that CI is still based on relationships with the people in traffic who provide information, but now do so with mobile communications technology. Yet, today there exist much better and various opportunities for participating in and gathering collective intelligence globally. The location of the participants is not important anymore, creating an opportunity for a greater number of individuals to participate in the facilitation of CI, and with a stronger reach.

Online Image

According to Internet World Stats (2016), the number of online users is over 3 and a half billion. The Internet is important as a medium because of its potential reach, but also as a source of diverse information and perspectives. The Nielsen Global Trust in Advertising (2015) indicates that consumers mostly have trust in recommendations from people they know (83%), branded websites (70%) and consumer opinions posted online (66%). On the other hand, consumers have less confidence in ads on TV (63%), brand sponsorships (61%), TV program product placements (55%) and other traditional forms of advertising (Nielsen, 2015). We can say that user reviews are seen as trustworthy information because they seem unbiased. The general thought is that they do not have any other interest to praise or criticize a certain product or organization than their interest in sharing their experience and satisfaction. With respect to these results, we can say that consumer online opinions represent sources of information about services and products that some users rely upon. Therefore, they are an important part of marketing analysis, but also valuable for consumer public relations. Clearly, the boundaries between marketing and PR are fading (Scott, 2007; Solis and Breakenridge, 2009; Brown, 2011). For example, some new promotion techniques have appeared in marketing, such as content marketing[1], which is much closer in form to the tasks assigned to the public relations profession. Seeing that PR is oriented toward managing communication with the public and relationship building, it is even harder to imagine successful public relations without considering user reviews.

[1] "Content marketing is the creation of quality branded editorial content across all media channels and platforms to deliver engaging relationships with all stakeholders, consumers and customers" (CIPR, 2014).

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

Measuring public relations' effectiveness is a challenging issue and an object of research for many PR experts (Macnamara, 2006; Lindenmann, 2006; Watson and Noble, 2005). For PR the goal is to influence or to change opinions, attitudes, and behavior of the public, which is the reason why effectiveness is mostly presented through the results on influence. Furthermore, the image of the organization being promoted by PR is the most visible and perceivable part credited to public relations. The term "image" is usually discussed in public relations handbooks (Tench and Yeomans, 2009; Theaker, 2001) and is a part of PR research. Don W. Stacks defines image research as a "research program or campaign that systematically studies people's perceptions toward an organization, individual, product, or service; sometimes referred to as a 'reputation study'" (2011: 335). Scott M. Cutlip *et al.*, (1985) point out that except for creating a positive organizational image, public relations can have the task to analyze public opinion in the market. In general, we can say that image is a "public view of something or someone" (Danesi, 2009: 154). But, if we specifically observe an organizational image, "it is the reflection of an organization in the eyes and minds of its publics" (Löwensberg, 2009: 239). Some of the authors see image as a multidimensional concept which may represent the outcome of public relations activities, but at the same time they are very critical towards using the term "image". For example, James E. Grunig (1993) thinks that the term image is associated with negative connotations such as creating illusions, imitating reality and embellishing someone or something. Furthermore, that is why the term "image" is often replaced with the term "reputation." In this study, image is interpreted as a short-term outcome of organizational activities that can be identified through individual attitudes and opinions. At the basic level we can say that attitudes can be determined as positive, negative or neutral; therefore determining sentiment in individual attitude, manifesting through positive, negative and neutral emotion, is the first step in understanding organizational image. Sentiment analysis represents a measuring tool for determining a positive or negative image (Jiménez-Zafra *et al.*, 2016; Soroka *et al.*, 2015).

It seems that what has been unavailable or hidden in the minds of target audiences reveals itself when we approach cyberspace. It is easier to express opinions when they are anonymous and in an apparently secure and known environment where users comment, and review products and services. On the other hand, users may expose themselves more because of their lack of awareness as to what extent their opinions are used and for what purposes. To one extent user comments are used for the idea of collective intelligence. From one perspective image can be facilitated from personal experience (for example, the experience of eating in a restaurant, watching a movie, or flying in an airplane). On the other hand, it can be a mediated experience, as it is the case with published online reviews, comments or news. It is the personal experience of some other individuals, on which users rely. Online organizational image is mostly dependent and characterized by user-generated content, but also by mainstream online media. Don W. Stacks (2006: 5) observes cyber image analysis as the measurement of Internet content published in chat rooms or discussion groups in cyberspace, referring to a client, product, or topic; it represents the measurement of a client's image everywhere online. Consumers are evolving as the markets evolve. They are becoming more demanding because they have more choices than ever before. Organizations are focusing more on advantages

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

which differentiate them from their competitors. Furthermore, today consumers have a new role, they have become "prosumers" (Toffler, 1980). Christofer Pihl (2013, in Ko *et al*., 2013) describes term "prosumers" as a combination of producers and consumers because consumers are now involved in the design and manufacture of products, they have much more influence in different areas of production, product development and distribution, which before was reserved only for employees of companies. If we observe consumers from this perspective, they participate as equally as employees in the role of organizational consultants. Users in the same way "manufacture" tourist destinations – accommodation, restaurants/gastronomy, "things to do" and attractions they offer. More importantly, they manufacture a destinations' image "as a continuous mental process by which one holds a set of impressions, emotional thoughts, beliefs, and prejudices regarding a destination due to information obtained from different channels" (Kim and Chen, 2015: 1). The assumption is that image has an effect on the overall success of a certain destination, and therefore marketing and public relations experts are prepared to listen what users have to say. From the perspective of place marketing, Eli Avraham and Eran Ketter (2008) emphasize that product marketing, the marketing of services in the private sector and social-public marketing are the most important set of tools. Restaurants and other services in the private sector, with their own image, take part in the destination image. They are image units that can be gathered in the collective intelligence process and create some other macro image unit.

MATERIAL AND METHODS

Research Questions and Hypotheses

Considering the idea of collective intelligence and the concept of online image in the context of "average" restaurants and sentiment analysis as a measurement tool, two research questions are examined:

RQ1: What is the online image of "average" restaurants on TripAdvisor?

RQ2: How accurate are computer sentiment analysis programs?

Two hypotheses were tested regarding previous research questions:

H1: The overall user rating "average", on a word scale, for the restaurants on TripAdvisor is different from the overall online image detected through sentiment in the user comments.

H2: Computer sentiment analysis can achieve accuracy of more than 60% for positive and negative sentiment estimations, in a comparison to referential human sentiment analysis.

Data Collection

The data for analysis contains user comments/reviews (N=786) from TripAdvisor (www.tripadvisor.com) – the travel website that recently became very important for researchers (Ayeh *et al*., 2013; Jeacle and Carter, 2011; O'Connor, 2010). TripAdvisor is known as the largest and the most popular travel site, which publishes reviews for hotels, vacation

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

rentals, attractions, restaurants and also allows commenting on forums. This research includes an analysis of user comments/reviews for restaurants in Dubrovnik, Croatia.

Dubrovnik is located on the Southern Adriatic coast and it is one of the most popular tourist destinations in Croatia, known as "Pearl of the Adriatic" and recognized by the famous Old City. The Old City of Dubrovnik has been under UNESCO's protection as a World Heritage Site since 1979 (Unesco.org, 2016). Being so popular, it is no wonder that Dubrovnik had over 85,500 reviews on TripAdvisor in December 2014. The Dubrovnik Tourist Board declared that in 2015 Dubrovnik had more than 3,300,000 overnights and over 932,000 arrivals. The majority of visitors are foreign tourists mostly coming from Great Britain, USA, Germany, France, Spain, Australia, Italy, Finland and Sweden (Poslovniturizam.com, 2015).

TripAdvisor allows users to publish their comments/reviews and rate restaurants with the scale: "terrible", "poor", "average", "very good", "excellent". There are over three hundred restaurants in Dubrovnik (Tripadvisor.com, 2015). Considering that TripAdvisor is already meant for users' evaluation of tourist destinations and provides overall ratings, this research is oriented on restaurants in Dubrovnik which have a total rating of "average." The word "average" on TripAdvisor is used as an adjective for describing a restaurant's value. The "average" rating, opposite to other ratings, seems hard to define. According to the Oxford dictionary, the adjective "average" has a few meanings: 1. Constituting the result obtained by adding together several amounts and then dividing this total by the number of amounts; 2. Of the usual or ordinary amount, standard, level, or rate; 3. Having qualities that are seen as typical of a particular person, group, or thing; 4. Mediocre; not very good (Oxforddictionaries.com, 2014). In terms of image, "average" restaurants could then be ordinary, typical or mediocre and not very good. This research examines whether the image of "average" restaurants is mostly positive, negative or neutral as the assumption is that there is a discrepancy between users' overall restaurant ratings on the word scale and their comments/sentiments expressed towards a restaurant.

The sample consists of 786 extracted English user reviews which date from February 2009 to December 2014. A unit of analysis represents one user review. By the end of year 2014, there were seventeen restaurants (each having more than ten reviews) in Dubrovnik characterized as "average": *Bistro Teatar*; *Baracuda; Konoba Amoret; Arka; Konavoka; Defne; Skybar; Cafe Bar Irish Pub Karaka; Restaurant Gusti; Konoba Pjatanca; Konoba Longo; Buffet Zvonik; Konoba Uvala; Hotel Dubrovnik; Aquarius; Konoba Sciabecco* and *Gallus* (Tripadvisor.com, 2014).

Method

For Theresa Wilson *et al*., (2005: 347), sentiment analysis is the task of identifying positive and negative opinions, emotions, and evaluations. Sentiment analysis is also known for its method of opinion mining. It represents a method for analyzing opinions based on polarities of expressed emotions. Kin Meng Sam and Chris R. Chatwin (2013) use polarities of happiness and sadness. Nevertheless, polarity is mostly presented as a negative, positive, or neutral emotion. This computer based language processing

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

recently became a focus of interest for scientists (Jakopović and Mikelić Preradović, 2014; Mostafa, 2013; Liu, 2012; Pang and Lee, 2008; Pang and Lee 2004). From a public relations perspective, sentiment is an emotion that we find in individual opinions about an organization, person, product or service. It is quite common in PR practice to estimate image through negative and positive polarities. We can say that image reflects emotions towards an organization and it is equivalent to expressed opinion. In respect to that, sentiment analysis is applied to image analysis.

According to the aim of this research and the stated hypotheses, sentiment analysis tools SentiStrength[2] and Sentiment Analysis with Python NLTK Text Classification[3] were validated and used for measuring online image. SentiStrength is a sentiment analysis open-source program that is used to analyse text. It uses polarities of positive, negative, and neutral emotions, but it also measures the intensity of positive and negative emotions. It uses scales from 1 to 5 (1 – neutral; 5 – extremely positive) and from -1 to -5 (-1 – neutral; -5 – extremely negative). SentiStrength is usually used for analysis of short English social web texts (Thelwall, 2010; Thelwall *et al.*, 2012). SentiStrength can be optimized furthermore through the creation and implementation of training data (Thelwall *et al.*, 2012). This research tested an unsupervised[4] version of SentiStrength. The results of SentiStrength analysis on platforms *BBC Forums*, *Digg*, *MySpace*, *Runners World*, *Twitter*, *YouTube* and *All 6* show a similar correlation in the key accuracy test between unsupervised and supervised (optimized) software version (Thelwall *et al.*, 2012). Sentiment Analysis with Python NLTK Text Classification, an open-source sentiment analysis tool described in the literature (Perkins, 2010; Bird *et al.*, 2009), enables the input and analysis of English texts which are expressed through positive, negative sentiment, or as neutral. It determines negative and positive polarity on the scale from 0 to 1. The strength of polarities is described with decimal numbers. For example, 0.1 represents lowest positive or negative sentiment strength and 1 absolute positive or negative sentiment strength.

Human analysis of sentiment in texts is used as a referential norm because of the human capability to understand the broader context as well as irony, sarcasm, metaphors and idioms. Since sentiment analysis programs still have problems with detecting a variety of meanings (Glucksberg, 2011; Liu, 2012; Pang and Lee, 2008), which humans can easily determine because of a combination of experience and knowledge, human analysis is set as the most accurate in determining positive and negative, or neutral emotions – manifesting as a positive, negative or neutral online image of an organization.

This research is not focused on the validation of intensity of positive and negative sentiments. The intensity of positive and negative emotions should rely on a large number of human coders to obtain the trustworthy standard[5] based on an inter-coder agreement. Rather, this research is aimed for an estimation of coherence between human coding and sentiment analysis tools, and the alignment among image evaluations. Therefore,

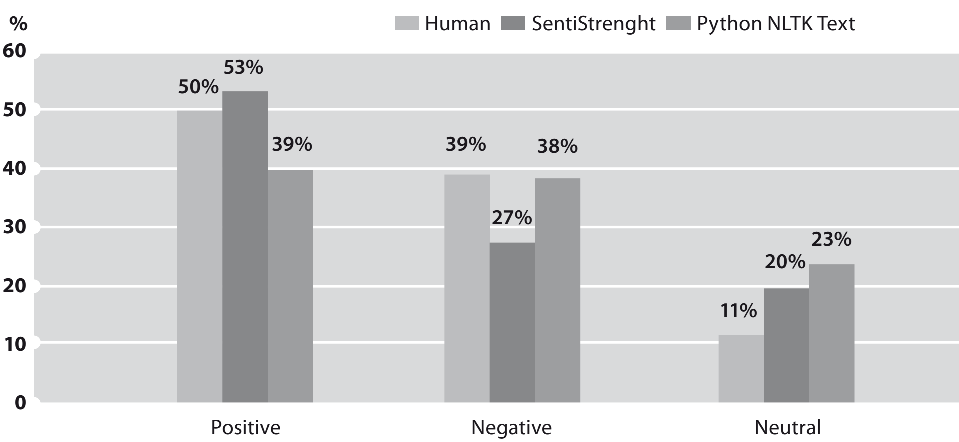
[2] http://sentistrength.wlv.ac.uk/ (15.07.2015).
[3] http://text-processing.com/ (15.08.2015).
[4] SentiStrength's lexicon can be optimized for a certain set of human-coded text. If SentiStrength is optimized, then it is in supervised mode; without optimization it is in the unsupervised mode (Thelwall, 2012).
[5] Gold standard

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

two independent coders conducted the analysis and inter-coder reliability was tested with Holsti's method on the sample of 50 randomly selected user comments/reviews for restaurants in Dubrovnik that were published on TripAdvisor. The result of inter-coder reliability test is 0,96.

RESULTS

The overall results concerning restaurants in Dubrovnik that were evaluated on TripAdvisor as "average," indicate that the ratio of positive, negative, and neutral sentiments in human analysis and SentiStrength and Python NLTK Text Classification analysis is similar (Graph 1). Positive emotions are mostly represented (Human analysis – 50%; SentiStrength – 53%; Python NLTK Text Classification – 39%). Negative emotions are second most represented (Human analysis – 39%; SentiStrength – 27%; Python NLTK Text Classification – 38%) and the percentage of neutral reviews is last by all (Human analysis – 11%; SentiStrength – 20%; Python NLTK Text Classification – 23%).



▲ *Graph 1*
Human and Computer Analysis of User Reviews

Considering the percentage of neutral reviews, it is possible to say that there are fewer neutral reviews in human analysis than in SentiStrength and Python NLTK Text Classification, because humans can better understand the context. The following example is evaluated as neutral by SentiStrength, but in human coding is estimated as low intensity negative emotion:

Very similar to the reviews that said: Don't go to restaurants with a lot of pictures in the menu. Service was ok but I felt not ok after the grilled squid. Choose it only if no other chances available.

The other example is estimated as neutral by Python NLTK Text Classification, however positive, with low intensity, in human coding:

Nice salads on a hot day, with a cool drink in the shade. Exactly what we needed. Friendly staff and decent food at reasonable prices.

The "average" restaurants are observed through a positive, negative or neutral image among users. The results show (Graph 1) that the overall rating "average" is just a rough estimation and so this analysis indicates that the real image remained hidden. The description of "average" restaurants is "ordinary," "typical" or "mediocre" and "not very good" (Oxforddictionaries.com, 2015), yet in Table 1 we can see that "average" restaurants with most comments/reviews have, in general, a positive image among users.

Table 1. **Percent Ratios of Positive, Negative and Neutral Restaurant Image**

	Positive image			Negative image			Neutral image		
	Human	**Senti**	**Python NLTK**	**Human**	**Senti**	**Python NLTK**	**Human**	**Senti**	**Python NLTK**
Restaurant									
Bistro Teatar, N=101									
	53%	58%	43%	38%	28%	30%	9%	14%	27%
Baracuda, N=72									
	53%	50%	46%	33%	33%	42%	14%	17%	12%
Konoba Amoret, N=72									
	57%	61%	46%	29%	20%	32%	14%	19%	22%
Arka, N=68									
	53%	53%	43%	32%	22%	35%	15%	25%	22%
Konavoka, N=66									
	53%	53%	44%	36%	20%	38%	11%	27%	18%

These are the examples of reviews that were evaluated as positive by all three analyses:

We love this nice restaurant! Ana Maria was a perfect waiter with good advice! We definitely recommend! Good food and very cozy location!
This is a really good spot, just off the main walkway in the Old Town. Service is very attentive and friendly, you aren't rushed when the meal is finished and the food, particularly the steaks, are excellent. Also good for people watching. We returned more than once. Thoroughly recommended!
This was inexpensive and the pizzas were very good. Set in a back street with plenty of atmosphere - recommended if you want something quick and tasty.

The following examples of user reviews were evaluated as negative by all three analyses:

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

- *Terrible Food, Terrible Service: went in for lunch on 25/10 had Spaghetti Milanese (Spaghetti Tomato Soup) and to say the food was poor would be an understatement.*
  Ordered seafood risotto. First brought me out a desert called risatta (my bf had main, why the hell would I order a desert for a main?) then my seafood risotto was brought out, it was rice with tomato soup from a can with four mussels. Pathetic. Don't recommend. Rude waiter was nice to Croatian customers and barely paid attention to us. Honestly would rather eat an instant pack of noodles.
  Very poor service even though the restaurant was half empty. Only seafood in seafood pasta was surimi!

Dan Jurafski et al., (2014) conducted research based on user reviews of restaurants on the website Yelp.com and their analysis showed that reviews for expensive restaurants and food are associated and described with metaphors of sexuality and sensuality, while fast food in cheap restaurants was related and described as "addicting" and "comfort food." Using the word frequency method 786 user reviews of Dubrovnik's "average" restaurants was analyzed. The frequency shows which positive sentiments occur most often: "good" (454 occurrences), "nice" (189 occurrences), "great" (167 occurrences), "like" (142 occurrences), "friendly" (135 occurrences), "excellent" (96 occurrences), "lovely" (76 occurrences). The most frequent negative sentiments are: "bad" (89 occurrences), "expensive" (78 occurrences), "poor" (64 occurrences), "avoid" (58 occurrences), "worst" (43 occurrences), "rude" (35 occurrences) and "terrible" (31 occurrences).

The accuracy of sentiment analysis tools is evaluated with Pearson's correlation. Correlation is firstly examined between human coding (HUMAN) and SentiStrength (SENTISTR) for 786 units of analysis (Table 2). The analysis shows the existence of a strong positive correlation (Table 2) with a coefficient of 0.413 and significance at the 1% level, for SentiStrength and human coding. The interpretation of a correlation coefficient is based on a range of +0.40 to +0.70 for a strong positive correlation. As a supplement to the results, the research of Mike Thelwall et al., (2010) showed that SentiStrength can predict positive emotion with an accuracy of 60.6% and negative emotion with an accuracy of 72.8%.

Table 2. **Pearson's Correlation for Human Analysis and SentiStrength**

Pearson Correlation (N=786)	**Human analysis** Coefficient	Significance
SentiStrength	**0.413**	**p<0.01** (2-tailed)

On the other hand, Pearson's correlation for human coding (HUMAN) and Python NLTK Text Classification (NLTKTEXT) for 786 units of analysis resulted with a coefficient of 0,199 and significance at the 1% level (Table 3). This represents a weak positive correlation of human coding and Python NLTK Text Classification. The interpretation of a correlation coefficient is based on a range of +0.05 to +0.20, representing a weak positive correlation. Python NLTK Text Classification matching with human analysis is lower than is the case with human coding and SentiStrength. Therefore we can say that overall Python NLTK Text Classification sentiment analysis is less accurate than SentiStrength analysis for estimating sentiment in user comments/reviews.

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

Table 3. **Pearson's Correlation for Human Analysis and Python NLTK Text Classification**

Pearson Correlation (N=786)	**Human analysis** Coefficient	Significance
Python NLTK	**0.199**	**p<0.01** (2-tailed)

The measurement of relative error gives insight into the discrepancy between sentiment tools and referent human coding estimations. It displays how large the error is in relation to the referent value. The relative error of SentiStrength in relation to referent human analysis (Table 4) is highest in the example of restaurant *Bistro Teatar* (9.2%) and lowest for restaurant *Baracuda* (5.3%), when it comes to positive image. There are no errors in estimations of positive image for restaurants *Arka* and *Konavoka*. The highest relative error of SentiStrength is in the estimation of the negative image of restaurant *Konavoka* (45.8%) and lowest for restaurant *Bistro Teatar* (26.3%). There is no error in estimation of negative image for restaurant *Baracuda*.

Table 4. **Relative Error for SentiStrength in a Relation to Human Analysis**

Restaurant	**Positive image**	**Negative image**
Bistro Teatar, N=101	9.2%	26.3%
Baracuda, N=72	5.3%	0%
Konoba Amoret, N=72	7.3%	33.3 %
Arka, N=68	0%	31.8%
Konavoka, N=66	0 %	45.8 %

For Python NLTK Text Classification the relative error referring to positive image (Table 5) is highest for restaurant *Konoba Amoret* (19.5%) and lowest for restaurant *Baracuda* (13.1%). Python NLTK Text Classification has the highest relative error in the estimation of negative image in the case of restaurant *Bracuda* (25%), and lowest for restaurant *Konavoka* (4.2%).

Table 5. **Relative Error for Python NLTK Text Classification in a Relation to Human Analysis**

Restaurant	**Positive image**	**Negative image**
Bistro Teatar, N=101	18.5%	21%
Baracuda, N=72	13.1%	25%
Konoba Amoret, N=72	19.5%	9.5%
Arka, N=68	19.4%	9.1%
Konavoka, N=66	17.1 %	4.2%

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

In comparison of these two programs (Table 4 and Table 5) we can say that SentiStrength is more accurate in analyzing a positive image on TripAdvisor, whilst Python NLTK Text Classification sentiment analysis provides better accuracy for a negative image.

DISCUSSION

Sentiment analysis as an approach for measuring emotions classified as positive, negative or neutral and manifested as a positive, negative or neutral image could be an applicable instrument for the evaluation of user comment/reviews and detecting organization's online image. We could say, by referring to the Oxford dictionary and the meaning of the term "average," that sentiment towards "average" restaurants should be more negative or neutral. The results indicate that, in the case of Dubrovnik restaurants, the overall user rating "average" is not a precise estimation of online image because the sentiment analysis of user comments revealed that every of these "average" restaurant has a mostly positive image. This confirmed the hypothesis that the overall user rating "average," on a word scale, for the restaurants on TripAdvisor is different from the overall online image detected through sentiment in the user comments (H1). These results give an insight on the existence of serious gaps between user ratings, based on scales available on the websites as imposed estimations, and user comments, as their true attitudes and opinions towards certain restaurant. This could be seen as an interesting phenomena which shows that ratings can be deceiving while user comments provide more fulfilled and trustworthy information. Furthermore, a consistency issue is present in the overall user reviews on websites. Average restaurants could be less consistent than higher ranked restaurants and for that reason they are marked as "average." Future research could be time-based to determine whether there are oscillations regarding the time of the meal (for example, dinner is better or worse than lunch) or day of the week (for example, meals on Monday are better or worse than on Friday). The analysis of word frequency also indicates that in user comments/reviews the same words with positive emotion, "good", "nice", "great", "friendly", "excellent", occurred more, while words that carry negative emotion are scattered. This displays that users more disparately express negative emotions and often use familiar phrases for expressing positive emotions. Analyzing word frequency in user reviews is useful for determining consumer needs.

Furthermore, the sentiment analysis programs SentiStrength and Python NLTK Text Classification were validated in comparison to human analysis in analysis of user comments/reviews for Dubrovnik restaurants which were rated as "average." One limitation of the study is that some other software for sentiment analysis is not tested. The results indicate that the sentiment analysis program SentiStrength proved to be more accurate considering the strong positive correlation in relation to referent human sentiment analysis. The matching between SentiStrength and human analysis is significant and it displays SentiStrength's applicability. However, Python NLTK Text Classification showed itself to be more accurate in analyzing negative sentiment or the negative image of restaurants, while SentiStrength is better at estimating positive emotions and positive

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

images of restaurants on TripAdvisor. Considering these results, the SentiStrength tool proved to be useful for estimating sentiment, especially when dealing with large datasets. The relative error showed considerable variations between sentiment analysis programs estimation of positive and negative restaurants' images. Nevertheless, the correlations tests and measurements of relative error enable confirmation of the hypothesis (H2). Therefore, we can say that computer sentiment analysis can achieve a precision of more than 60% in a relation to referential human sentiment analysis.

CONCLUSION

The idea of collective intelligence (CI) that we recognize through user interaction and user-generated content represents a good example of CI in practice. In this study the idea is described and determined by referring to previous findings in the area of collective intelligence. As users interact and publish their comments/reviews, they are facilitating CI but are also involved in the creation of online image. Image is an important part of the public relations profession as it defines how publics see organizations. Therefore, the efforts of public relations often reflect a positive or negative image. Considering the relation between collective intelligence and organizational image, we can conclude that CI facilitated through user comments/reviews is involved in the creation of online image. Online image is partially a product of collective intelligence and is above all determined with sentiment found in the attitudes and opinions of individuals towards an organization. With respect to that, we can say that emotions often remain hidden, but are more frequently expressed in cyberspace because of the imaginary anonymity and the feeling of security when using a computer in a usually comfortable environment, for example at home. As public relations efforts are oriented on influencing opinions and attitudes of various publics, the creation of positive image represents the most visible and essential part of the public relations profession. Whereas, we can say that online image could be measured with sentiment analysis tools. Still, sentiment analysis tools have difficulties in estimation because of idioms, metaphors, sarcasm, irony, spelling mistakes and the length of comments/reviews. The discrepancy in neutral sentiment estimations, in comparison of referent human sentiment analysis with sentiment tools, is a consequence of a program's inability to recognize the context as humans can, especially when user comments/reviews are based on low intensity positive or negative emotions. However, this study showed that in public relations research sentiment analysis should be taken as a possible method and instrument in the process of measurement and evaluation of PR activities. It is encouraging that PR practitioners could soon have instruments capable of accurately measuring organizational image. Thereby sentiment analysis tools could, with further development and enhancement, save them time for conducting research and evaluating effects of public relations.

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

References

>Aitamurto, Tanja (2011) The Impact of Crowdfunding on Journalism. *Journalism Practice.* 5 (4): 429-445. DOI: 10.1080/17512786.2010.551018.

>Alag, Satnam (2009) *Collective Intelligence in Action*. Greenwich: Manning Publications Co.

>Avraham, Eli and Ketter, Eran (2008) *Media Strategies for Marketing Places in Crisis*. Oxford: Elsevier.

>Ayeh, Julian K., Au, Norman and Law, Rob (2013) "Do We Believe in TripAdvisor?" Examining Credibility Perceptions and Online Travellers' Attitude toward Using User-Generated Content. *Journal of Travel Research* 54 (4): 437-452. doi: 10.1177/0047287512475217.

>Bird, Steven, Klein, Ewan and Loper, Edward (2009) *Natural Language Processing with Python*. Sebastopol: O'Reilly.

>Brown, Rob (2011) *Public Relations and the Social Web: How to use social media and Web 2.0 in Communications*. London: Kogan Page.

>CIPR (2014) *Content marketing*. http://www.cipr.co.uk/ (25.07.2016).

>Cooren, Francis (2004) The Communicative Achievement of Collective Minding: Analysis of Board Meeting Excerpts. *Management Communication Quarterly* 17(4): 517-551. DOI: 10.1177/0893318903262242.

>Cutlip Scott, M., Center, Allen H. and Broom, Glen M. (1985) *Effective public relations*. New Jersey: Prentice-Hall.

>Danesi, Marcel (2009) *Dictionary of Media and Communications*. New York and London: M.E. Sharpe.

>Glucksberg, Sam (2011) Understanding Metaphors: The Paradox of Unlike Things Compared. *Affective Computing and Sentiment Analysis* 45: 1-12. DOI: 10.1007/978-94-007-1757-2_1.

>Grunig, James E. (1993) Image and Substance: From Symbolic to Behavioral Relationships. *Public Relations Review* 19 (2): 121-139.

>Gunawardena, Charlotte N., Hermans, Mary Beth, Sanchez, Damien, Richmond, Carol, Bohley, Maribeth and Tuttle, Rebekah (2009) A theoretical framework for building online communities of practice with social networking tools. *Educational Media International* 46 (1): 3-16. DOI: 0.1080/09523980802588626.

>Internet World Stats (2016) *Internet Users in the World*. http://www.internetworldstats.com/stats.htm (24.07.2016).

>Jakopović, Hrvoje and Mikelić Preradović, Nives (2014) Mining Web and Social Networks for Consumer Attitudes towards Government-Owned Croatian National Airline. *International Journal of Computers* 8: 128-135.

>Jeacle, Ingrid and Carter, Chris (2011) In TripAdvisor we trust: Rankings, calculative regimes and abstract systems. *Accounting, Organizations and Society* 36 (4-5): 293–309. DOI: 10.1016/j.aos.211.04.002.

>Jenkins, Henry (2002) *Interactive audiences? The collective intelligence of media fans*. http://labweb.education.wisc.edu/steinkuehler/curric606/readings/Jenkins2002.pdf (10.03.2015).

>Jiménez-Zafra, Salud M., Martín-Valdivia, M. Teresa, Martínez-Cámara, Eugenio and Ureña-López, L. Alfonso (2015) Combining resources to improve unsupervised sentiment analysis at aspect-level. *Journal of Information Science* 42 (2): 213-229.

>Jurafsky, Dan, Chahuneau, Victor, Routledge, Bryan R. and Smith, Noah A. (2014) Narrative framing of consumer sentiment in online restaurant reviews. *First Monday* 19 (4). http://journals.uic.edu/ojs/index.php/fm/article/view/4944/3863 (12.03.2015). DOI: 10.5210/fm.v19i4.4944

>Kim, Hyangmi and Chen, Joseph S. (2015) Destination image formation process: A holistic model. *Journal of Vacation Marketing* 22 (2): 1–13. DOI: 10.1177/1356766715591870.

>Ko, Eunju, Chun, Eunha, Song, Sangah and Kim, Kyung Hoon (2013) Which content types increase participation in fashion social platforms?. *Journal of Global Scholars of Marketing Science: Bridging Asia and the World* 23 (3): 297-313. DOI: 10.1080/21639159.2013.79350.

>Lee, BongKoo, Choong-Ki, Lee and Lee, Jaeseok (2014) Dynamic Nature of Destination Image and Influence of Tourist Overall Satisfaction on Image Modification. *Journal of Travel Research* 53 (2): 239–251. DOI: 10.1177/0047287513496466.

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

>Leimeister, Jan Marco (2010) Collective Intelligence. *Business & Information Systems Engineering* 2 (4): 245-248. DOI: 10.1007/s12599-010-0114-8.
>Lévy, Jean-Pierre (1997) *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Cambridge, Massachusetts: Perseus Books.
>Lindenmann, Walter K. (2006) *Public Relations Research for Planning and Evaluation*. Florida: Institute for Public Relations.
>Liu, Bing (2012) *Sentiment Analysis and Opinion mining*. San Rafael: Morgan & Claypool Publishers.
>Livingstone, Randall M. (2016) Models for Understanding Collective Intelligence on Wikipedia. *Social Science Computer Review* 34 (4): 497-508. DOI: 10.1177/0894439315591136.
>Löwensberg, Danny (2009) Corporate image, reputation and identity, pp. 237-251 in Tench, Ralph and Yeomans, Liz (eds) *Exploring Public Relations*. Harlow: Pearson Education Limited.
>Macnamara, Jim (2006) *PR METRICS – Research for Planning & Evaluation of PR & Corporate Communication*. http://www.mediasearch.co.nz/documents/PR%20Metrics%202006.pdf. (12.02.2015).
>Malone, Thomas W. (2008) What is collective intelligence and what will we do about it?, pp. 1-4 in Tovey, Mark (ed.) *Collective intelligence: creating a prosperous world at peace*. Oakton, VA: Earth Intelligence Network.
>Mostafa, Mohamed M. (2013) An emotional polarity analysis of consumers' airline service tweets. *Social Network Analysis and Mining* 3 (3): 635-649. DOI: 10.1007/s13278-013-0111-2.
>Nielsen (2013) *Global Trust in Advertising*. http://www.nielsen.com/us/en/insights/reports/2015/global-trust-in-advertising-2015.html (20.05.2016).
>O'Connor, Peter (2010) Managing a Hotel's Image on TripAdvisor. *Journal of Hospitality Marketing & Management* 19 (7): 754-772. DOI: 10.1080/19368623.2010.508007.
>Pang, Bo and Lee, Lillian (2008) Opinion Mining and Sentiment Analysis. *Information Retrieval* 2 (1-2): 1-135. DOI: 10.1561/1500000011.
>Pang, Bo and Lee, Lillian (2004) A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. *ACL '04 Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*. http://aclweb.org/anthology/P/P04/P04-1035.pdf?CFID=247293914&CFTOKEN=70536811 (25.03.2015). DOI: 10.3115/1218955.1218990.
>Perkins, Jacob (2010) *Python Text Processing with NLTK 2.0 Cookbook*. Birmingham: Packt Publishing.
>Peters, Michael A. and Reveley, James (2015) Noosphere rising Internet-based collective intelligence, creative labour, and social production. *Thesis Eleven* 130 (1): 3-21. DOI: 10.1177/0725513615575932.
>Poslovniturizam.com (2016) *Dubrovnik u 2015. godini postigao rekordne turističke rezultate (Dubrovnik in 2015 achieved record tourism results)* .http://www.poslovniturizam.com/brojke-trendovi/dubrovnik-u-2015-godini-postigao-rekordne-turisticke-rezultate/1963 (20.07.2016).
>Sam, Kim Meng and Chatwin, Chris R. (2013) Ontology-Based Sentiment Analysis Model of Customer Reviews for Electronic Products. *International Journal of e-Education, e-Business, e-Management and e-Learning* 3 (6): 477-482. DOI: 10.7763/IJEEEE.2013.V3.282.
>Solis, Brian and Breakenridge, Deirdre K. (2009) *Putting the Public Back in Public Relations*. New Jersey: Pearson Education, Inc.
>Soroka, Stuart, Young, Lori and Balmas, Meital (2015) Bad News or Mad News? Sentiment Scoring of Negativity, Fear, and Anger in News Content. *The ANNALS of the American Academy of Political and Social Science* 659 (1): 108-121. DOI: 10.1177/0002716215569217.
>Stacks, Don W. (2011) *Primer of Public Relations Research*. New York: Guilford Press.
>Stacks, Don W. (2006) *Dictionary of Public Relations Measurement and Research*. Florida: Institute for Public Relations.
>Surowiecki, James (2004) *The Wisdom of Crowds*. New York: Anchor Books.
>Tench, Ralph and Yeomans, Liz (2009) *Exploring Public Relations*. Harlow: Pearson Education Limited.
>Theaker, Alison (2001) *The Public Relations Handbook*. London and New York: Routledge.

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

>Thelwall, Mike, Buckley, Kevin and Paltoglou, Georgios (2012) Sentiment strength detection for the social Web. *Journal of the American Society for Information Science and Technology* 63 (1): 163-173. DOI: 10.1002/asi.21662.

>Thelwall, Mike, Buckley, Kevin, Paltoglou, Georgios, Cai, Di and Kappas, Arvid (2010) Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology* 61 (12): 2544–2558. DOI: 10.1002/asi.21416.

>Toffler, Alvin (1980) *The Third Wave*. New York: Bantam Books.

>UNESCO (2016) *Old City of Dubrovnik*. http://whc.unesco.org/en/list/95 (20.06.2016).

>Viktorsson, Cano (2013) Traffic Radio as a Precursor to Smart Travel Planning Systems: The Challenge of Organizing "Collective Intelligence". *Journal of Urban Technology* 20 (4): 43-55. DOI: 10.1080/10630732.2013.863439.

>Watson, Tom and Noble, Paul (2005) *Evaluating Public Relations*. London: Kogan Page.

>Wilson, Theresa, Wiebe, Janyce and Hoffmann, Paul (2005) Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*. http://acl.ldc.upenn.edu/H/H05/H05-1044.pdf (10.02.2015). DOI: 10.3115/1220575.1220619.

>Woolley, Anita Williams, Aggarwal, Ishani and Malone, Thomas W. (2015) Collective Intelligence and Group Performance. *Current Directions in Psychological Science* 24 (6): 420-424. DOI: 10.1177/0963721415599543.

IZVORNI ZNANSTVENI RAD / DOI: 10.20901/ms.7.13.6 / PRIMLJENO: 28.05.2015.

RAZOTKRIVANJE *ONLINE* IMIDŽA „PROSJEČNIH" RESTORANA NA TRIPADVISORU

Hrvoje Jakopović

SAŽETAK *Kolektivna inteligencija može se interpretirati kao djelovanja pojedinaca iz kojih proizlaze kolektivni učinci. Što je više komentara korisnika o određenoj temi u* online *prostoru, to je veća mogućnost da će se određeni stav nametnuti kao dominantni okvir kroz koji se promatra ta tema. To ujedno pruža uvid u vrijednost korisničkih komentara, recenzija i ocjena za odnose s javnošću. Može se zamijetiti kako mnogo neodlučnih pojedinaca prilikom donošenja odluke o kupnji određenog proizvoda ili usluge poseže za informacijama drugih korisnika koji su s tim proizvodom ili uslugom već imali iskustva. Takve informacije imaju učinak na donošenje odluka i na poticanje ponašanja. U slučaju odnosa s javnošću, kolektivna inteligencija producirana kroz komentare i recenzije korisnika uključena je u proces stvaranja imidža. Ovo istraživanje mjeri* online *imidž na temelju ideje o kolektivnoj inteligenciji uz pomoć analize sentimenta koja se opisuje kroz upotrebu aplikacija za obradu prirodnog jezika te se time dobiva uvid u korisničke stavove i mišljenja. Autor analizira korisničke komentare/recenzije za restorane u Dubrovniku koji su ocijenjeni kao „prosječni" na internetskoj stranici TripAdvisor. Efikasnost računalne analize sentimenta uspoređena je s referentnom ljudskom analizom sentimenta. Rezultati pokazuju u kojoj mjeri alati za analizu sentimenta mogu biti primjenjivi instrumenti za procjenu pozitivnog, negativnog i neutralnog sentimenta i razotkrivanje* online *imidža.*

KLJUČNE RIJEČI

KOLEKTIVNA INTELIGENCIJA, *ONLINE* IMIDŽ, ANALIZA SENTIMENTA, KORISNIČKI KOMENTARI, ODNOSI S JAVNOŠĆU, TRIPADVISOR

Bilješka o autoru

Hrvoje Jakopović :: Sveučilište u Zagrebu, Fakultet političkih znanosti, Zagreb **::** hrvoje.jakopovic@fpzg.hr