ISSN 2063-6415
www. metszetek.unideb.hu

Abstracts

New ways in exporting Society
The potential of donation.based digital data collection

KMETTY, ZOLTÁN

More and more digital data is being generated every day, and more and more social science analyses are using Twitter, Instagram, or Facebook data. Many international and national studies have already explored the social science opportunities and dilemmas raised by the phenomenon of ‚big data' – but the issue of ‚access to data' has only been touched upon tangentially. And access to data is becoming increasingly difficult. What can we do if market players close the access to their data, and, if we find data available, the Research Ethics Board tells us to stop? The answer is simple: go to the users and ask them for the data. This approach is what the literature calls data donation. This paper will describe the data donation approach in detail, focusing on how researchers can access data through users on the current major Western platforms. The practical feasibility of data donation access will be illustrated based on a domestic pilot study.

KEYWORDS: data collection, data donation, social media, big data, Facebook

The challenges of supervised machine learning in sociological applications

NÉMETH, RENÁTA

The sociological applications of supervised machine learning, already well proven in industrial/ business applications, raise specific questions. The reason for this specificity is that in these applications, the algorithm is tasked with learning complex concepts (e.g. whether a tweet contains hate speech). Supervised learning consists of learning to classify previously annotated (hate speech/non-hate speech) texts by the algorithm, looking for characteristic text patterns. The questions that arise are: how to prepare annotation? How can a hermeneutic challenge such as hate speech recognition be performed by annotators? Are routinely applied, detailed annotation guidelines helpful? The article also discusses how large companies perform coding on crowdsourcing platforms, and describes AI bias, which in this case means that annotators themselves introduce bias into the data. I illustrate these issues with our own research experiences.

KEYWORDS: supervised machine learning, annotation, crowsourcing, AI bias

ISSN 2063-6415
www. metszetek.unideb.hu

Hungarian Videoblogger Networks Online

HEGEDŰS DÁNIEL

The web 2.0 phenomenon and social media – without question – not only reshaped our everyday experiences, but they have established an environment for new types of social practices and social actors. The demotization (Turner 2010) effect of such technologies has created entirely new fields where celebrities might emerge from: one of them is videoblogging. Many video bloggers gained great reputation through peculiar micro-celebrity practices (Marwick 2015, Senft 2012), and, as a result, became key figures in distributing ideas, values and knowledge in today's society. These cognitive patterns are disseminated with a discursive apparatus that is largely based on social media activity, including posts, tweets, self-imagery and the videos themselves, which are tied to a certain logic according to environmental affordances, creating the possibility for fans to interact (share, comment, like, retweet etc.) with artifacts of the celebrity. This mechanism puts the celebrity in a so-called expert system (Giddens 1990) position as they provide adequate schemas of attitude, mentality or behavior. Most importantly, all of these public interactions are accessible for scholars to conduct scientific research. With the help of the SentiOne application this research attempts to reconstruct online networks of video bloggers based on mentions, which either occurred in an artifact (post, video description etc.) or in a fan comment. Apart from the network itself, SentiOne enables us to get insights regarding each individual connection established in it with different types of aggregated data.

KEYWORDS: digital sociology, expert systems, networks, discourse analysis

Empirical analysis of the judgment of unconditional basic income through YouTube comments

OLÁH, ESZTER

One of the world's largest video-sharing platforms is YouTube, where viewers can comment on the videos and their topics. The aim of this study is to examine the values and opinions about unconditional basic income according to the comment sections of several Youtube's videos which topic is the previously mentioned UBI which is receiving increasing attention in parallel with today's economic and social changes. Our research works with a mixed method, data collection, storage, sentiment analysis and the bag of words method which were implemented using IT procedures, while categorization was done through manual coding. The results of the sentiment analysis show that positive arguments appear to a lesser extent in the comments. Positive arguments have value characteristics such as inclusion, the principle of the right to exist, justice and freedom. Among the positive arguments feasibility enjoys the highest support. Negative category values arise more frequently, so the emphasis on the values of injustice, exclusion, unaffordability, and performance-orientation is dominant in the analyzed comments.

KEYWORDS: text mining, computer text analysis, sentiment analysis, unconditional basic income, work-based society

ISSN 2063-6415
www. metszetek.unideb.hu

What's the matter? A text mining analysis of political topics and user engagement on politicians' Facebook pages during the 2018 Hungarian general election campaign

BENE, MÁRTON

The research investigates the way users interact with leading topics of the 2018 Hungarian general election campaign on candidates' Facebook pages. It expects that the prominent (immigration, corruption) and campaign-related topics generate more user engagement, while policy topics and mobilization content are less interacted. It also tests the theory of issue ownership in relation with user engagement. These expectations are tested on a dataset that includes all posts (38030 posts) posted by all candidates during the campaign (511 candidates). Topics are identified by text mining methods. The study demonstrates that corruption, development policy and campaign are highly engaged topics, while immigration was more interacted only on opposition politicians' pages since the followers of pro-government candidates engage less with immigration-related content. The most surprising result is that a reversed issue ownership effect can be detected since politicians are generally less successful with their own topics.

KEYWORDS: social media, issue ownership, immigration, user engagement, text mining

Where to go, net generation? Lifestyle-based segments of the Hungarian youth

SÜTŐ, ANNA

My study attempts to explore the lifestyle-based segments of the Hungarian youth through an innovative methodology based on social media data, incorporating the dimension of digitization into the creation of lifestyle groups. The examination of the segments' lifestyle attitudes is assisted by a review of the related theoretical milieu approaches, international and Hungarian empirical milieu researches.

KEYWORDS: lifestyle, milieu, youth research, Facebook

Social media communication in the digital medical space

SIMON, SÁRA

In the environment of 21st century technology, the transformation of information acquisition of health care and patients has had an increasing emphasis. Despite the earlier authoritative doctor-patient relationship, a need for an equal, cooperation-based communication has emerged and there are so many digital healthcare projects to achieve this (Koskova 2015).

Information acquisition on the internet has allowed patients that based on the increasingly available medical information they acquire information about their condition, become part of patient communities, ask for second opinions, and become committed helpers of their doctors in their disease (Meskó et. al 2017).

This can be especially true for patients with rare diseases, where a diagnosis might take even a decade, the patient needs lifelong condition maintenance and treatment, if it is available. While the proportion of patients with rare diseases is low compared to the whole of society, the number of such patients is approximately 30 million in Europe (EURORDIS), which means patients and

ISSN 2063-6415

their relatives need not only a harmonized health care system, but extensive information so that they can live with the rare disease with less difficulty.

The aim of our study was to present the options of information acquisition in the social media, focusing on Twitter, via an interdisciplinary and social approach. In this study therefore we carried out a Big Data based social media analysis based on #Asthma and #CysticFibrosis databases of the Symplur corporation. This study results contain the complete online communication of 7 years (2012–2019) regarding these hashtags. The analysis has few levels including semantic research, stakeholder and hashtag review, engagement, and the whole tweet activity exploration.

KEYWORDS: digital healthcare, e-patients, Big Data, cystic fibrosis, asthma

Classification of depression-related online forums using Natural Language Processing

MÁTÉ, FANNI

The study of the phenomenon of depression is not new in sociology, but since the depression is becoming a wider social problem, it is still a relevant issue today. In addition to the bio-medical and psychological aspects of depression, the sociological perspective is becoming more noteworthy in the discourse about the causes of depression. In the research of the discourse on depression, the online texts offer many new possibilities, as the forum's anonymity and accessability make the online seeking for help popular. In this research, natural language processing (logistic regression) was applied to find patterns in the definition of depression in lay discourses. These methods make it possible to analyze a large amount of text – which would have been difficult to process with human resources. During the analysis, 67 857 posts of English-speaking online forums were categorized along the categories of the scientific discourse about depression. This study presents the first results, which shows logistic regression classifier performs like the annotators. . Although the research has analyzed English-speaking forums, my findings may be useful to anyone observing abstract sociological concepts in online texts written by users.

KEYWORDS: depression, Natural Language Processing mental health, online forums