

BIG DATA

Digital Marketing and Trendwatching

Edited by
Agnieszka Dytman-Stasieńko
& Agnieszka Węglińska



BIG DATA

Digital Marketing and Trendwatching

BIG DATA

Digital Marketing and Trendwatching

edited by
Agnieszka Dytman-Stasieńko
& Agnieszka Węglińska

Reviewers:

Dr. hab. Marcin Brocki (Jagiellonian University in Krakow)

Dr. hab. Andżelika Kuźniar (SGH Warsaw School of Economics)

Proofreading:

Monika Budzyń-Kelly

Cover design and DTP:

Bartosz Harlender



The publication of the book is funded by the project – Big Data, Digital Marketing and Trendwatching – a new master degree program taught in English at the University of Lower Silesia co-financed by the European Social Fund under the Operational Program Knowledge Education Development. Project number M034/16.

© Copyright by National Center for Research and Development, Wrocław 2018

ISBN: 978-83-7977-341-1



Oficyna Wydawnicza ATUT – Wrocławskie Wydawnictwo Oświatowe
50-439 Wrocław, ul. Kościuszki 142, tel. 71 342 20 56
wydawnictwo@atutowyczna.pl; www.atut.ig.pl

Contents

Introduction	7
Michał Jacuński	
Measuring and Analysis of Digital Marketing	11
Bartłomiej Łódzki	
Essential Internet Tools for Media Studies Research.....	28
Marta Majorek	
Darknet and Medical Big Data. Deep Internet as a Space for Illegal Trade in Medical Information	48
Agnieszka Dytman-Stasieńko	
(Big) Data Activism – Trends, Challenges and Opportunities ...	65
Ivan Valchanov	
Data Journalism. Essence, Technology, Development	83
Michał Kuś	
Data Journalism Projects in Polish Mainstream Media Landscape. Case Study of Biqdata and Polityka Insight	93
Maria Nikolova	
Crisis PR and Media: PR Tactics for Communication with Journalists and Media.....	107
Marek Zimnak	
Marketing Our Project Abroad	124
Agnieszka Węglińska	
Internet Astroturfing and Security Threats: The Protests in Defence of Poland's Courts, Bots and Disinformation.....	140
Sylwia Siekierka	
Where Does Creative Class Meet Trendsetters?.....	159

Introduction

According to the most recent report of the International Data Corporation, commissioned by Seagate Technology, by 2025 the amount of data in the global datasphere will have increased to 163 ZB (trillion gigabytes), or ten times more than was measured at the end of 2016. In its report, IDC analyses the trends that are driving the global increase in data, and the consequences of such a rapidly growing datasphere. It enumerates five key trends that will intensify the role of data in the world:

- 1) **The evolution of data from business background to life-critical**
– previously inaccessible or little-used data will become essential, not only for society, but also in individual lives. IDC estimates that by 2025, 20% of data will be critical in daily life, and 10% of that, hypercritical.
- 2) **Embedded systems and the Internet of Things (IoT)** – by 2025, the average individual will work with connected devices almost 4,800 times a day (one interaction every 18 seconds), generating a huge amount of data.
- 3) **Mobile and real-time data** – by 2025, over a quarter of data will be real time in nature, of which 95% will be data in the Internet of Things. Furthermore, data will need to meet the criteria of instant availability.
- 4) **Cognitive/artificial intelligence (AI) systems that change the landscape** – cognitive systems (machine learning, natural language processing, and artificial intelligence), the development of which will drive the growth of data, can increase the frequency, elasticity, and immediacy of the analysis of data in many areas.

- 5) **Security as a critical foundation** – the report's authors have noted a significant gap between the amount of data being produced now that must be secured, and the amount that actually is secure. According to the authors, in 2025, 90% of all data will need to be secure to a certain extent, but not even half will be secured (Reinsel, Gantz, Rydning 2017).

The IDC report confirms the growing and critical role of data both in the global sphere, and in the lives of communities and individuals; thus, it can be assumed that data education is essential. To pursue this goal, the University of Lower Silesia (ULS) in Wrocław (Poland) now offers *Big Data, digital marketing and trendwatching*, a new graduate programme, taught in the English language. Its innovative curriculum combines knowledge from the fields of economics, communications, digital marketing, public relations, statistics, and new technologies – essential in work involving the gathering, analysis, interpretation and ultimately visualisation of data. Education in the skills that will be needed in the future presents a challenge to our institution. Programmes of study must be adapted to the needs of the job market, and teaching methods must meet market demands. It is essential to identify and deploy skills that will be needed at work. Training programmes in the field of social communication should deal with elements of theory, technical problems and social and interpersonal issues; such an education, recommended in the documents of the European Union, is also demanded by employers (Frączek 2014, McKinsey 2016).

An important part of the ULS's programme consists of practical education about software tools dedicated to personal branding, social and mobile marketing, viral and buzz marketing, content management and creation, and video marketing. Students are also provided with skills related to trendwatching and trendsetting. These include searching for regular patterns in data, identifying dependencies and predicting trends, and using these skills for effective marketing in the digital world. These are key skills for the future in the digital world of the datasphere. According to the Digital Economy Lab's report, *Tendencje na rynku – 2017 (Market tendencies, 2017)*, it is not enough for one to upgrade professional knowledge to hold one's own in the dynamically changing job market; the development of digital and interpersonal skills, and flexibility to deal with such change, are also essential (DEL 2017).

This volume presents selected aspects of the issues raised above.

It opens with an article by Michał Jacuński, who characterises the phenomenon of Internet marketing, distinguishing it from traditional marketing and demonstrating the benefits of access to data. The author notes the volatility of the Internet marketing industry, which is dependent on constantly developing technology.

In conducting research about the media, one must be acquainted with tools for collecting, processing and interpreting data. These are presented in a further article by Bartłomiej Łódzki.

Marta Majorek examines the dark side of Big Data and the Darknet: the theft and illegal trade of sensitive data, in this case medical information. In this context, she analyses two opposed tendencies, involving the limits of freedom and the right to anonymity.

Agnieszka Dytman-Stasieńko analyses Big Data from the perspective not of big players such as governments and corporations, but of those working for social change. Data activism has become a very important part of their work, confronting them with a number of new challenges while offering new opportunities.

In turn, Ivan Valchanov analyses data journalism, describing its essence and the technology that it makes use of. He also notes the emergence of data journalism as a professional activity.

Michał Kuś focuses on data journalism as well, in this case taking Poland as an example. He analyses the ways in which the development of this type of journalism has influenced the implementation of data journalism tools in the largest media organisations, and the development of the practice of data journalism practices.

Marek Zimnak analyses aspects of the internationalisation of business activity, noting that this involves various branches of the social sciences, with a particular emphasis on knowledge of marketing and its tools.

Maria Nikolova describes public relations tactics that may be employed in dealing with the media and journalists in crisis situations. She presents the social media challenges that the public relations expert faces when managing a crisis in the era of “fake news”.

Agnieszka Węglińska analyses the phenomenon of astroturfing, drawing on the example of a campaign in defence of the Polish courts in July 2017. Astroturfing is unequivocally deplored as a form of manipu-

lation. The author points to the need to verify sources and to be aware of the possibility that data has been manipulated, and reality distorted, for political ends.

Sylwia Siekierka focuses on the interdisciplinary nature of research into the creative industries. An important paradigm of the creative economy is the immateriality of such factors as ideas, talent, and pleasure, connecting them to phenomena such as trendsetting. Siekerka points out the interaction between market tendencies and lifestyle.

The texts in this volume deal with both theoretical and practical aspects of Big Data, digital marketing and trendwatching. This field of research, into an area that is experiencing unusually dynamic development, demands continual attention. It also calls upon both future practitioners and regular network users to take it upon themselves to deepen their education. We hope that the presented volume will help bring this about.

References:

- Digital Economy Lab (2017), *Aktywni+ Przyszłość rynku pracy*, http://docplayer.pl/47624255-Aktywni-przyszlosc-rynku-pracy-raport-gumtree-2017-aktywni-przyszlosc-rynku-pracy.html#show_full_text, accessed 4.01.2018.
- Frączek M. (2014), *Rynek pracy Unii Europejskiej – stan, perspektywy i główne determinanty rozwojowe*, In: M. Makuch (ed.) “Współczesny rynek pracy. Zatrudnienie i bezrobocie w XXI wieku”, http://www.bibliotekacyfrowa.pl/Content/64012/Wspolczesny_rynek_pracy.pdf, accessed 4.01.2018.
- Gantz J., Reinsel D., Rydning J. (2017), *Data Age 2025: The Evolution of Data to Life-Critical Don't Focus on Big Data; Focus on the Data That's Big*, An IDC White Paper, SEAGATE, <https://www.seagate.com/files/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf>, accessed 4.01.2018.
- McKinsey (2016), *Where Machines Could Replace Humans and Where They Can't*, http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/_where-machines-could-replace-humans-and-where-they-cant-yet, accessed 4.01.2018.

Measuring and Analysis of Digital Marketing

Abstract: The aim of the article is to discuss forms of measurement and analytics, which serve to achieve marketing goals through better online performance. Digital marketing offers a range of opportunities, which enable preparation and implementation of digital marketing strategy of products and services. However, it operates in a constantly changing environment, which requires sophisticated tools and new approaches to understand customer journey.

Keywords: digital marketing, marketing efficiency, data-driven marketing

Introduction

Digital marketing can be defined in many ways, however, in most cases it refers to traditional understanding of marketing role in selling products. Financial Times *Lexicon* (2018) describes it as *the marketing of products or services using digital channels to reach consumers*. The key objective is to promote brands through various forms of digital media. Digital marketing can be yet another term to describe the meaning of electronic marketing (e-marketing) or Internet marketing or web marketing, which refers to the management and execution of marketing using electronic media and wireless media in conjunction with digital (big) data about customers behaviour and characteristics (Chaffey *et al*, 2009, 10). It has to be stressed that currently digital marketing extends beyond Internet marketing to include channels that do not require the use of the Internet. The Institute of Direct and Digital Marketing (IDM 2018) brings more detailed explanation of what digital marketing involves: 1) applying technologies which form online channels to market; 2) creating and supporting of marketing activities aimed at achieving

profitable acquisition and retention of customers; 3) recognizing the strategic importance of digital technologies. It must be however remembered, that digital marketing does not occur in isolation, in many cases it is integrated with other activities, which brings us to the concept of multichannel marketing or integrated marketing communications. Such an approach includes the combination of all aspects of marketing communication, like advertising, sales promotion, public relation, direct marketing, personal selling and so on.

Digital marketing offers a range of opportunities for established multichannel organisations, which enable preparation and implementation of digital marketing strategy of products and services across purchase cycle in many more ways than in traditional marketing (Chaffey 2012, 14-15): advertising medium, direct-response medium, platform for sales transactions, lead generation method, distribution channel, customer service mechanism, relationship-building medium.

There is a noticeable difference in the approach adopted in contemporary marketing, especially when collating distribution channels and media types, and also the particular media used in traditional and digital marketing (Table 1).

Table 1. Traditional vs digital marketing

Traditional marketing		Digital marketing	
Television	<ul style="list-style-type: none"> · advertising medium · distribution channel · weak relationship-building · limited platform for sales transactions 	Social media	· advertising medium
Radio		Mobile devices	· direct-response medium
Newspapers, dailies		Paid placement	· platform for sales transactions
Magazines		SEO (Search Engine Optimization)	· lead generation method
Outdoor		Emails, newsletters	· distribution channel
Direct mailing		Online video	· customer service mechanism
Telemarketing		Paid search listings	· relationship-building medium
Leaflets		Display advertising: rich media, text ads, static images etc.	

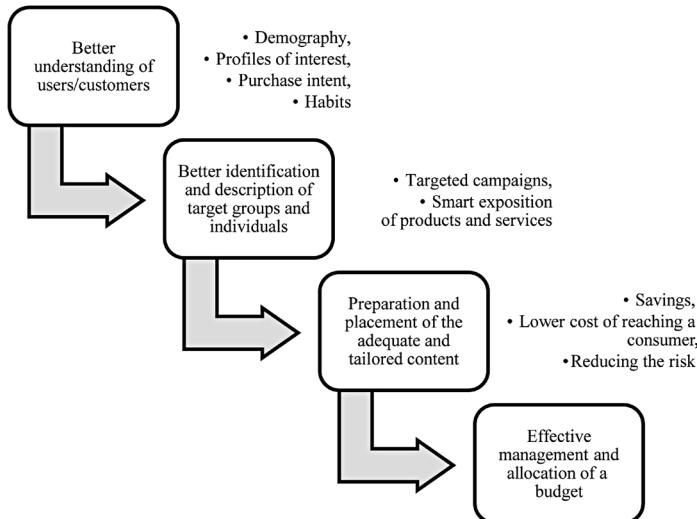
Source: Own elaboration.

Błażewicz (2018, 14), an experienced Polish marketer and successful entrepreneur, considers that in a shift from traditional to digital marketing (new) marketers can switch from campaign-based actions to process-based actions, which builds more durable company value. According to another enthusiastic marketer (Oza 2017), a list of advantages of digital marketing over traditional marketing is enormously plentiful. They primarily include: low cost, huge return on investment (ROI), ease of measurement, ease of adjustment, brand development, ease of sharing, precise targeting, global reach, segmentation, greater engagement.

In its operations, digital marketing employs, gathers and processes Big Data coming from different sources. Through proper data refining and ordering, it is possible to reach customers more accurately while lowering costs, and to better manage the marketing budget, especially advertising spending. Data-driven marketing is becoming a significant part of digital marketing, while Big Data in the marketing context is termed *data management platform* (DMP). Such platforms aggregate audience data from various online and offline sources: web analytics tools, mobile web, mobile apps, e-commerce portals, behavioural and demographic data, CRM data, POS data, social networks, online video, Smart TV data etc. Advanced quantitative data analysis techniques for data-driven marketing include (Li 2013): frequency distribution, descriptive statistics, statistical testing, cross tabulation, establishing correlations, linear regression, text analytics, automated content analysis.

Many researchers and analysts confirm that Big Data, and thus DMP, is based on the principle of four V's (IBM 2018): *volume* (big scale of data), *variety* (diversity of data, different forms of data), *velocity* (high speed of data, significant dynamics of streaming data), *veracity* (uncertainty of data). Mróz (2016, 130) rightfully argues this is a resource of great utility to marketers but it also brings a challenge both in terms of acquiring and analysing historical data, but also with respect to forecasting and prediction.

Fig. 1. Benefits from data-driven marketing



Source: Own elaboration based on the report *Konsumpcja treści online a marketing*, www.iab.org.pl/wp-content/uploads/2016/04/Raport_Konsumpcja_tresci_online_20161.pdf, accessed 06.02.2018, p. 54-55 (IAB, 2016).

Framework and context of digital marketing related research

Kozielski notices (2016, 83-84) that marketing measuring is part of research into company operational efficiency, but despite the relatively long-lasting interest in the matter, marketing measuring still has not developed measuring standards, and the source literature has only been attempting to systematize the field for the last few years. According to the researcher, marketing measuring is possible on several levels and in several fields: on the level of management (the strategic field), on the level of strategic business units (the systems and market processes field), e.g. the owner, and finally on the level of operational activity (the operations and tools field), the level of managers' activity.

Digital technologies are rapidly changing environment within which companies operate, thus Kannan and Li (2016, 25-31) developed and described a concise framework for research in digital marketing that

highlights the touchpoints in the marketing process as well as in the marketing strategy process. In the proposed framework they provided an overview of research issues of five areas: 1) *consumer behaviour*, which reflects studies on customer journey through different stages of the buying process across digital and non-digital environments, 2) *social media* and *user-generated-content* (UGC), which tap into the study of structured data and the unstructured textual content, including semantic analysis and the role of influencers, 3) *platforms* and *two-sided markets*, where independent third-party entities connect buyers, sellers, companies, the crowd creating various social and business networks, 4) *search engines*, which bring together three key players: the search engine, the advertiser/company and the customer and examine relationship between them and the performance metrics, 5) *contextual interactions*, impacted by geography and location, regulations on privacy and regulations against the piracy of content.

In the Table 2 there are listed detailed research developments, which are having and will have a significant impact on digital marketing.

Table 2. Digital technologies and marketing environment: research issues

Area of focus	Research developments
Consumer behaviour	<ul style="list-style-type: none"> a. Stages of buying process, purchase funnel, and impact of digital environments and digital devices b. Information acquisition, search, information processing and decision aids in digital environments c. Buyer behaviour across digital and non-digital environments d. Customer trust and risk perceptions in digital environments
Social media and UGC	<ul style="list-style-type: none"> a. Electronic Word-of-Mouth (eWOM) and motivation for eWOM b. Dynamics in eWOM posts and their impact on sales c. How eWOM posts influence other posts? d. Social networks, identification and targeting of influencers e. eWOM and fake reviews

Area of focus	Research developments
Platforms and two-sided markets	<ul style="list-style-type: none"> a. Network effects in online platforms, information asymmetry and impact on sales b. Impact of competition on two-sided content platforms c. Issues in crowdsourcing and using platforms for innovations
Search engines	<ul style="list-style-type: none"> a. How should search engines price and rank keywords? b. How should advertisers choose specific keywords and bid on them? c. Relationship between rank, click-through rate and conversion rate, and decision support for optimal bidding d. Synergy between organic search and paid search
Contextual interactions	<ul style="list-style-type: none"> a. Interaction between geography/location and digital environments b. Impact of regulatory environment – Privacy concerns and effectiveness of digital marketing c. Impact of piracy of content

Source: Kannan P.K., Li "Alice" H. (2017), *Digital Marketing: A Framework, Review and Research Agenda*, "International Journal of Research in Marketing", Volume 34, Issue 1, March 2017, p. 25.

Online marketplace analysis is yet another form of assessing digital marketing strategy. According to Chaffey (2014), it involves review of four main areas and summarising the dynamics of the interactions between:

- a) potential online audience – what are the profiles, characteristics, behaviours, needs and wants of audiences from digital media platforms?;
- b) customers – a similar analysis to potential audience, but also gaining feedback on their views of your current online presence. Defining personas is an essential technique here;
- c) competitors – benchmarking the capabilities of direct competitors for your services and indirect competitors for traffic or review of other sites for “out-of-sector” best practice;
- d) online intermediaries, influencers and partners – establishing the strengths and weaknesses of intermediaries who can include current or potential partners to reach and influence your audience.

These include different types of publishers from general online news to more specific blogs and affiliates. Intermediaries also include social networks and communities.

It appears that, unlike professional marketers, customers do not differentiate their experiences based on the diverse media environments and touch points; instead, they can only state their loyalty and confidence in a brand, product or medium. The changes in digital marketing have far-reaching consequences, which do not always meet with understanding in the professional world. Firstly, it can be observed that it is the customers who – often publicly – generate more content and messages than those managing the brands; secondly, content coming from marketing creators is becoming more and more liquid; and thirdly, to a degree, customers are becoming co-owners of the brand. The above developments have an impact on ways of measuring the effectiveness of digital marketing operations, which can be conducted on three levels:

- 1) the level of the consumer – by obtaining statistical data – mainly socio-demographic – from user profiles; by analysing motivation and consumer behaviour, which allows to establish key consumer values; by studying user experience (UX), including the mode of using tools and applications, and consuming particular contents etc.,
- 2) the level of the brand – based on the findings relating to brand image and the stage reached in the product life cycle,
- 3) the level of business development – defined as the measuring of purchase intention as well as the actual purchasing actions and decisions undertaken by consumers.

Stephen (2016) points out that an emerging theme in recent years is how digital and social media environments, often stipulated by digital marketing, impact consumer behaviour. He mentions two main, but different, perspectives. On one hand, it can be thought of as *environment-integral*, so that digital environments influence consumer behaviour in those environments, and on the other *environment-incidental*, which presumes that digital environments influence behaviour in other, unrelated environments. Therefore, he claims, it is an interesting question, to learn how various characteristics of digital environment, like the degree of exposition to reviews, opinions, decisions and choices made through-

out social media and web, can impact subsequent customers behaviours.

Since the early 2000s, digital marketing has been changing and constantly there occur many new trends, of which most require new or constantly improving analytical solutions. According to Forbes Agency Council (Forbes 2017) the top fifteen trends of 2018 include:

- augmented reality integrated through social media,
- collapse of the influencer market,
- understanding of customer journey,
- personalization along all stages of the sales funnel,
- more targeted ads with specific purpose,
- native advertising,
- professional live video,
- video marketing,
- conversational user interfaces,
- contextual targeting strategies,
- cost per experiment as a leading metric,
- digital integrated into offline experiences,
- voice marketing,
- predictive algorithms.

Measurement and analytics as part of marketing strategy

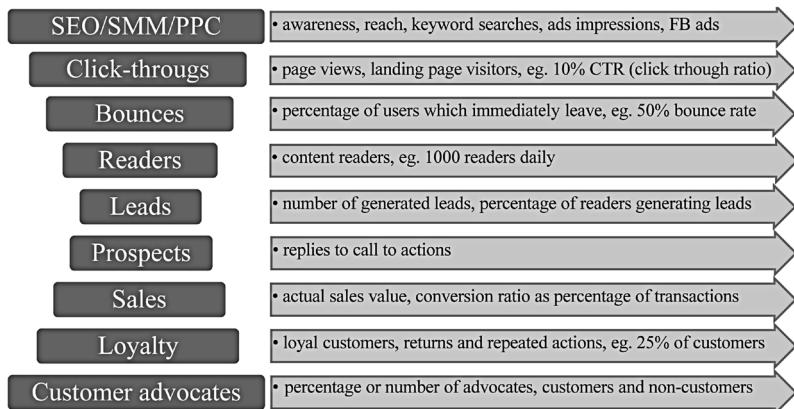
In principle, measurement and analytics serve to achieve the following goals in the field of better marketing performance: to increase CTR for ads and social shares, to reduce bounce-rates, to increase opt-in frequency, to increase email opening and response rates, to increase customer purchase rates, to increase frequency and value of repeat purchases, to increase number of advocates.

In marketing communication, the concept of the marketing or sales funnel has gained popularity (Fig. 2), and it can also be applied to the measuring of operational effectiveness.

At the top of digital marketing funnel the visitors/clients first encounter your online brand. It is when visitors arrive at the website or landing page or social network page etc. through the following actions: search engine optimization (SEO), social media marketing (SMM), pay per click (PPC) advertising. If they are interested parties and are will-

ing to establish relationship, they potentially become content readers, which can transform into actions from the middle of a funnel, like leads generation or interaction. Then, at the bottom of the funnel there occurs the opportunity to convert attracted individual into a customer.

Fig. 2. Marketing funnel: analytics and measurement



Source: Own elaboration based on CoolerInsights, <http://coolerinsights.com/2016/09/how-to-optimize-your-digital-marketing-funnel/>, accessed 04.02.2018.

When it comes to evaluating whether the influence of digital marketing on consumer decisions is long- or short-term, the answer is provided by research conducted by the consulting company McKinsey (McKinsey & Company, 2016): it established that 10-30% of sales is attributable to current influence, which means that almost 70% of decisions are made due to the long-term influence of marketing.

What impacts the effectiveness of digital marketing operations? Some key and universal features of good marketing practice, which can be evaluated, include: regular publication of content, language accuracy, clear message, high quality of visual/video content, variation of content (text, video, graphics etc.), reaching the desired target group, using the media-mix and rich media, and the ability to respond/interact/ engage in dialogue in appropriately short time. The proper employment of these elements translates into the ability to maintain a good ratio of mention/discussion sentiment, whose positive and neutral rate should significantly outweigh the negative one.

The measurability framework consists of easy to measure and hard to measure indicators and it is important that they are exercised with different frequency, e.g. weekly, monthly, quarterly, annually. The more metrics are added into analysis, the more effort is required to understand the relevance of a single factor. The worlds of digital marketing and public relations attempted to adopt metrics taxonomy which, primarily driven by social media and networks, could become an applicable model of performance assessment in various media types. The known PESO model consists of paid, shared, owned and earned media and was evaluated by Bartholomew (Metricsman, 2018), who created a matrix (Table 3) which integrates media types and five performance dimensions: exposure, engagement, influence, impact, advocacy. Let us explain further key elements of this measurement concept.

By *exposure* I assume the degree to which a company's target group or an individual customer is exposed to the communication about its products or services. As de Vreese and Neijens (2016, 69) stated, the measurement of how people are exposed to media content, which is crucial for the understanding of various online media use and effects, has been a challenge for a long time. Online media landscape, in which individuals are exposed to a diversity of messages anytime, anywhere, and from a great variety of sources on an increasing number of different media platforms, has complicated the measurement of exposure even more. Nevertheless, digital landscape offers new possibilities to map exposure and to measure the impact of marketing goals in various types of media.

Influence on awareness is the capacity or power of persons or things to be a compelling force on brand awareness, which most often represents the level of consciousness of a company. It shows a potential customer's ability to not only recognize an overall brand image, but also to associate it with a desired company's product or service and its attributes.

Impact on decisions is an effect of an activity within community or target group as well as a direct effect on individuals, which ends up with transactions or decisions made. Awareness efficiency help building the right audience, that is group of people who know you exist. Engaged people take actions to support business or cause and it is important to analyse their communication patterns. Those who declare interest gave permis-

sion to communicate and reach them and digital marketing's role is to optimize it for conversions. Those who are already committed and take actions require analysis of behaviours, like transactions, donations etc.

Engagement measures the number of interactions for an online business' social media efforts. Engagement has been a common metric for evaluating online media performance but does not necessarily translate to sales increase. Engagement encapsulates a variety of actions, such as a sharing, liking, mentioning etc. It is any interaction a fan/follower/ customer has with brand's or corporate social media content that shows they are interested in your post. This includes both positive and negative aspect of engagement, because one cannot *a priori* assume that every form of engagement supports brand or business. Low or high engagement rate can be calculated by dividing the overall engagement volume by the number of users that could have triggered the action.

Online *advocacy* is a process of supporting ideas or people and enabling web users to express views, or to review and to recommend items. It can include many online activities or collective actions and campaigns in media which aim at decision making process.

Table 3. Measurement of marketing goals across various types of online media

Media type / Goals	Exposure	Engagement
Paid (refers to all third-party forms of paid content, mostly advertisement)	Impressions Reach Frequency Viewability GRPs TRPs Number of new fans (fans growth etc.)	Increasing the engagement Click-throughs Page views Interactions (replay, enter data etc.)
Owned (any kind of web properties, like a website, blog or social media channel, that is under control and is unique to a brand)	Unique visitors Visits Payment and organic coverage	Return visits Page views Interactions Subscriptions Links Subscriptions of newsletters
Shared (online participation and interaction across channels and technologies controlled by consumers around shared branded and non-branded content)	Organic impressions Organic reach Number of followers Number of posts	Reactions in social media: Likes Comments Shares Replies Retweets Mentions
Earned (essentially online word of mouth or content picked up by third-party sites, in order to write about clients and their products and services)	Number of posts (publications), including headings Impressions Message delivery	Hashtag usage Mentions Contest entries/participants Personal data left
All, convergent media	Views, including live views	

Source: Own elaboration based on Metricsman blog, <https://metricsman.wordpress.com/>, accessed 31.01.2018.

Influence on awareness	Impact on decisions	Advocacy
Awareness Purchase consideration Purchase intent Likelihood to recommend Brand attributes Changing the sentiment (settings) to the company / brand	Website visits Attend event Sales conversion Download coupon Leads captured (opt-ins) Promo redemptions Declared intent of purchase / employment of etc.	Placing information in a specified group / community
Considerations Tell a friend Brand attributes and equities Changing awareness	Sales Leads Information requests Downloads Cost savings Number of products / tickets sold.	Mentions among opinion leaders, influencers, publications
Positive quotes, evaluations Considerations Tell a friend Brand attributes and equities Changing awareness	Visit store Attend the event Sales Vote for an issue Satisfaction Loyalty Downloads	Third-party endorsement Recommendations rate
Awareness Consideration Purchase intent Associations with issues/topics	Visit website Attend event Download coupon Leads captured Promo redemptions Declared intent of purchase / employment of etc.	Mentions in earned channels
	Submissions, Applications, Requests	Reviews Ratings Recommendations

Finally, it is worth mentioning other important aims of digital marketing, which, although not directly connected to sales operations, relate nevertheless to market presence and the ability to survive on it. These include *employer branding*, connected with staff recruitment and references provided by a respected employer; the influence on attitudes and opinions of Internet users who are not customers might also be crucial; the interest from influencers; comprehensive *competitor analysis*. Thomas (2018) considers that the *competitor analysis* (CA) could form a key part in creating a digital marketing strategy. A full digital marketing CA should be a comprehensive document that encompasses all areas of digital marketing, including: websites, search rankings, organic and paid search, online presence, online reputation, site analytics capability. If a company finds a competitor doing something significantly better it can try to choose to emulate their strategy. If on the contrary weaknesses were noticed, one should focus efforts on these areas to easily take advantage of an 'easy win'.

Conclusions

With the hybridization of media (Chadwick 2014) and the growing and increasingly heterogenic public, it is becoming more and more challenging methodologically- and technically-wise to accurately measure the effectiveness of marketing operations. Measuring the impact of digital marketing operations on customer and stakeholder relations has implications for brand and product development, as well as for the financial well-being of market contributors. Following results and marketing optimization – under the conditions of the growing importance of interactive communication modes – is necessitated by the changes taking place among customers and in the marketing environment. It is evidenced by the ratio of digital marketing spending to general marketing spending in companies. According to estimates by Forrester Research (Forrester 2014), the expenditure of companies in USA to cover these costs will rise from 24% in 2014 to 35% in 2019. The biggest chunk here goes to cover the costs of search marketing, display advertising, with less spending on social media and email marketing. The above situation results from the shift in the time spent consuming different types of media. The data can

be successfully applied to the Polish market, where – according to different estimates – current expenditure on online marketing exceeds 20% of marketing budgets. It is expected that in the next two or three years, both in Poland and in developed countries, digital media will replace traditional television as the most consumed media. The implications to be expected are, on the one hand, marketing automation using the potential of Big Data; on the other hand, the need to seek new and better automated modes of analysing and measuring marketing operations, using technological solutions which, although designed by man, employ machine learning mechanisms, virtual and augmented reality (VR and AR), artificial intelligence (AI), and advanced semantic analysis allowing to translate machine language into natural language. This leads to the conclusion that the challenges faced by researchers and marketing practitioners centre around: measuring and the ability to reliably analyse the effectiveness of operations, the integration of operations carried out using different channels, the constant need to adjust technological and IT solutions to the requirements of the marketing sector, and finally the search for measurable effects of digital marketing, resulting from relations with stakeholders and non-customers.

References:

- Błażewicz G. (2018), *Marketing Automation Revolution Using the Potential of Big Data*, Warszawa, PWN.
- Chadwick A. (2013), *The Hybrid Media System. Politics and Power*, Oxford: Oxford University Press.
- Chaffey D. (ed.) (2009), *Internet Marketing: Strategy, Implementation and Practice*, Essex: Pearson Education.
- Chaffey D., Ellis-Chadwick F. (2012), *Digital Marketing: Strategy, Implementation and Practice*, Essex: Pearson Education.
- Chaffey D. (2014), *Online marketplace analysis for digital marketing*, <https://www.smartinsights.com/digital-marketing-strategy/online-marketplace-analysis/>, posted on 29th April 2014, accessed 08.02.2018.
- CoolerInsights, *How to Optimize Your Digital Marketing Funnel*, <http://coolerinsights.com/2016/09/how-to-optimize-your-digital-marketing-funnel/>, posted on 28th September 2016, accessed 04.02.2018.

Financial Times *Lexicon*, digital marketing definition, <http://lexicon.ft.com/Term?term=digital-marketing>, accessed 05.02.2018.

Forbes (2017), *How Digital Marketing Will Change In 2018: 15 Top Trends*, <https://www.forbes.com/sites/forbesagencycouncil/2017/12/18/how-digital-marketing-will-change-in-2018-15-top-trends/>, posted on 18th December 2017, accessed 06.02.2018.

Forrester Research Digital Marketing Forecasts, 2014 to 2019 (US) (2014), <https://markstaton.files.wordpress.com/2015/12/forrester-digital-marketing-forecast-2014-2019.pdf>, posted on March 2014, accessed 30.01.2018.

IAB, Raport 2016, *Konsumpcja treści online a marketing*, www.iab.org.pl/wp-content/uploads/2016/04/Raport_Konsumpcja_tresci_online_20161.pdf, Związek Pracodawców Branży Internetowej IAB Polska, posted on 5th April 2016, accessed 06.02.2018.

IBM Big Data and Analytics Hub (no date), <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>, accessed 14.02.2018.

McKinsey & Company (2016), *Discussions on Digital: Measuring Marketing's Impact*, <https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/discussions-on-digital-measuring-marketings-impact>, posted as a podcast on March 2016, accessed 01.02.2018.

Kannan P.K., Li "Alice" H. (2017), *Digital Marketing: A Framework, Review and Research Agenda*, "International Journal of Research in Marketing", Volume 34, Issue 1, March 2017, pp. 22-45, <https://doi.org/10.1016/j.ijresmar.2016.11.006>, Elsevier.

Kozielski R. (2016), *Pomiar działań marketingowych – proces i obszary / Measuring Marketing – Process and Dimensions*, "Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu/Research Papers of Wrocław University of Economics", 2016, vol. 460, p. 82-89, DOI: 10.15611/pn.2016.460.08, accessed 07.02.2018.

The Institute of Direct and Digital Marketing (IDM), <https://www.theidm.com/>, accessed 04.02.2018.

Li J. (2013), *Quantitative Data Analysis Techniques for Data-Driven Marketing*, <http://www.iacquire.com/blog/quantitative-data-analysis-techniques-for-data-driven-marketing-2>, posted on 12th April 2013, accessed 05.02.2018.

Metricsman blog, *Social Media Metrics and Measurement Continue to Evolve*, <http://metricsman.wordpress.com>, posted on 9th June 2014, accessed 31.01.2018.

Mróz B. (2016), *New Information Technologies: Implications for Business Strategies and Marketing Communication*, in: G. Mazurek, J. Tkaczyk (eds.), "The Impact of the Digital World on Management and Marketing", Poltext, Warszawa 2016, pp. 125-140.

Oza H. (2017), *10 Advantages of Digital Marketing Over Traditional Marketing*, <https://prowly.com/magazine/2017/10/11/advantages-of-digital-marketing-over-traditional-marketing/>, posted on 11th October 2017, accessed 07.02.2018.

Stephen A.T. (2016), *The Role of Digital and Social Media Marketing in Consumer Behavior*, “Current Opinion in Psychology”, vol. 10. pp. 17-21, <http://dx.doi.org/10.1016/j.copsyc.2015.10.016>.

Thomas R. (2018), *Four Types of Analysis Critical to Your Digital Marketing Strategy*, MyCustomer web magazine, <https://www.mycustomer.com/marketing/strategy/four-types-of-analysis-critical-to-your-digital-marketing-strategy>, posted on 2nd February 2018, accessed 09.02.2018.

Vreese C de., Neijens P. (2016), *Measuring Media Exposure in a Changing Communications Environment*, “Communication Methods and Measures”, Volume 10, 2016 – Issue 2-3, <https://doi.org/10.1080/19312458.2016.1150441>.

Essential Internet Tools for Media Studies Research¹

Abstract: The digitilisation and development of social media have led to changes in forms of communication. In today's digital environment, Internet users generate a huge amount of data. Media researchers face the challenge of collecting, processing and interpreting this information. It is becoming essential to use Internet analytical tools that support the research process. The aim of this article is to present selected scholarly and commercial applications. The author describes the range of their potential use, devoting particular attention to programmes that may be used without specialised technical or programming knowledge.

Keywords: media studies research, Internet research, analytical tools, social media, Big Data

Introduction

The development of social media has brought a range of long-term consequences with it. They may be examined at many levels. One of these is the ways in which various communication channels may be used, both by senders and receivers. The former, represented especially by the traditional media, use social networks to strengthen their position, especially to maintain their role as opinion makers. This is possible because a message may now reach an ever-wider audience, especially the younger

¹ This article broadens and complements a text that was originally published in 2016 as: "Lokalne – Regionalne – Transnarodowe. Rola mediów w kształtowaniu wspólnot", (Local, regional, trans-national: the role of media in shaping communities., in K. Wasilewski, K. Kopecka-Piech, eds., Gorzów Wielkopolski: Stowarzyszenie naukowe "Polska w Świecie", pp. 139-157.

generation for whom social media has become one of the most important sources of information. As P. Szews notes, *social media have evolved from a series of tools that facilitate communication, providing a source of entertainment, to media that now include practical, multi-function, multi-platform, usable tools that journalists and politicians use, and which have become an important source of information for typical users. Many of them treat these tools uncritically, not applying an objective filter, trusting in them without checking and confirming them with other sources* (2014, 96). A report by the Nielsen research institute has confirmed that as early as in 2012, 605 of Americans stated that they were better informed because of social media (Nielsen 2012). This figure is continually increasing. By 2017, 67% of Americans were turning to at least one social medium for their news. These observed changes include older and less-educated people (Sherarer, Gittfired 2017). The popularity of new media, and the changes in forms of communication that arise from it, has also determined developments in fields of study involving media research. Traditional research of media content and public opinion research, based on surveys, no longer suffice to observe and interpret the emerging changes; new technologies are helping. Along with analytical Internet tools that are appearing and so-called “digital methods”, media scholars have been exploiting the possibilities of conducting comparative studies in many areas, including those that are trans-national. Media researchers, without needing technical knowledge, have taken advantage of opportunities to analyse large data samples while simultaneously reducing the cost of such tasks. Well-chosen tools are allowing researchers not only to collect data, but to observe the reaction of Internet users in real time.

Digitalisation and Big Data

Digitalisation is taking place at several levels: in library collections, media archives and data bases that are available online (including the articles of the world’s largest publishers). As a result, media scholars are confronted with the most significant opportunity for development that they have had in years. These changes have been stimulated by new trends in the commercial market, and especially in the sector broadly

referred to as marketing and advertising. International concerns and interactive agencies have a great amount of data at their disposal. Those who can process this data, drawing suitable interpretations from it, can prepare more interesting offers for their current and future clients, and use their own resources more effectively. Media scholars should follow these trends and take them into account in their research. Work with large amounts of data is directly connected with the concept of Big Data. The term was first used in 2005 by R. Mougalas, of the O'Reilly Media agency (a year after the term "Web 2.0" was coined).² In his opinion, although the term has been understood somewhat differently in various fields, similar mechanisms underlie it. Above all, Big Data entails the collection, processing, analysis and visualisation of data. It also involves comparing data and drawing suitable conclusions from it. The concept usually refers to large data sets, virtually unmanageable using traditional tools (Halevi, Moed 2012, 3). T. Elliott has pointed out the development of the concept over time. At first, it was linked to the so-called "three Vs": Volume, Velocity and Variety. At present, however, Big Data is identified by eleven Vs, including Validity, Veracity, Value and Visibility (Elliott 2013). For analysts, data have become a reliable, high-quality and essential source of information. The humanistic and social sciences, including media studies, are among the latest that have been harnessing the research potential created by the application and processing of Big Data. From 1999 to 2004, in well-known academic journals such as *The American Sociological Review*, *The American Journal of Sociology* and *Social Forces*, only one article may be identified where data was collected with the use of Web-based research techniques. A decade later, the use of the potential of the Internet had risen sharply, owing to the efforts of researchers involved in such areas as communication about health matters, educational applications of the media and political communication. The last of these took its inspiration from examples of the commercial use of processed data by companies and other institutions conducting market studies. These billion-dollar industries have gradually won over representatives of the humanistic sciences. At the same time numerous and replicable analyses have helped strengthen the conviction that Big Data is a valuable tool for research.

² The first mention of "Big Data", understood as unimaginable amounts of data, dates from 1663, when John Graunt carried out a pioneering statistical analysis of data related to an epidemic that was then wasting Europe (Foote 2017).

tion that opinion research using Internet techniques, thus based on Big Data, generates representative data. Further, such research has led to a decrease in measurement errors and improved access to segments of the population that are not easily accessible. Active Internet research has influenced the development and upgrading of applied techniques. They have increased the usefulness and reliability of collected data (Farrell, Petersen 2010, 114). Because analytical tools are being used skilfully in academic environments, it is becoming possible to monitor the activities of commercial concerns, and to explain these processes. In the era of Big Data, the possibility of collecting data about citizens – voters – has become more common. An example of this is the observation of the reactions of citizens to information spread by political campaign staff or their public relations agencies. As T. Amerding has pointed out (2016), the collection of such information often takes place without the knowledge and control of its very subjects, who also have no influence on the safety of the collected data. Moreover, citizens themselves publish their own opinions with the help of social media, and in this way create a wide field of research and observation of the new digital reality. Analysis using Big Data allows for the monitoring of the most important political and media actors, as well as typical network users.

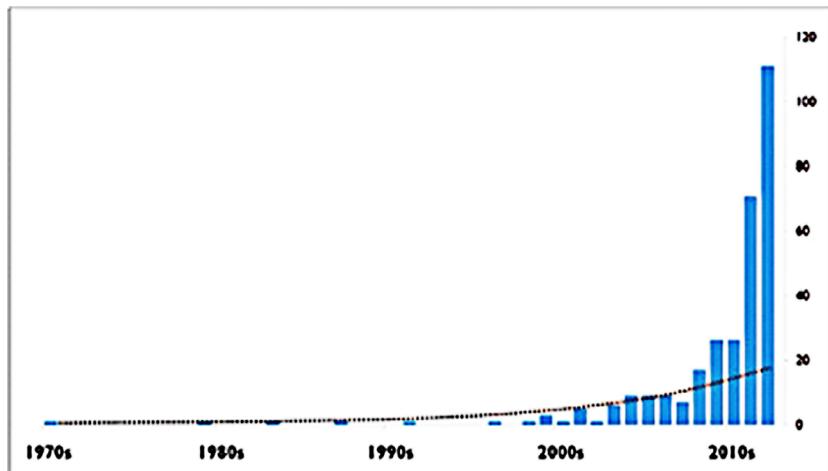
Literature review

As changes in the market and the appearance of new analytical Internet tools have appeared, the need for pertinent academic research has grown. There is already an increasingly rich and accessible literature on this subject. This is reflected, for example, in texts concerned with the problems involved in Big Data. The sharp increase in the number of these texts has coincided with the intensive development and use of new technology. At first – at the turn of the 1990s – most publications about the subject were from the exact sciences (information technology, engineering, mathematics, management, physics, astronomy and biotechnology). By the end of the first decade of the present century, the subject of Big Data was already being dealt with in the social sciences (Chart 1). Most materials have been published in the United States, China, Germany and Japan, countries where modern technology plays a

key role in the economy (Halevi, Moed 2012, 5). At present, the experiences of different disciplines are being noted in media studies literature. The research can be described as interdisciplinary, and it often has a trans-national and thus comparative dimension. Selected examples of such work, most closely related to the theme under discussion here, will be presented below.

Among the publications that deal with the shaping of international media studies research – the direction of its development, the research approaches taken, and its cognitive potential – is *International Media Research. A Critical Survey* (Corner et al., 2005). The authors analyze the issue from the perspectives of researchers not only from the Anglo-Saxon countries, but also from Central-Eastern Europe and Latin America.

Chart 1. Big Data as an academic research subject, over time



Source: Halevi and Moed, 2012, 4.

The authors of *Internationalizing Internet Studies* introduce the subject of current Internet research, presenting results from various parts of the world, especially from Asian countries (Goggin, McLelland 2009). The last twenty years have seen dramatic technological change. Forms of communication – including the ways that information is presented, used and disseminated – have changed under the influence of social media, mobile devices and robotics. This phenomenon and its conse-

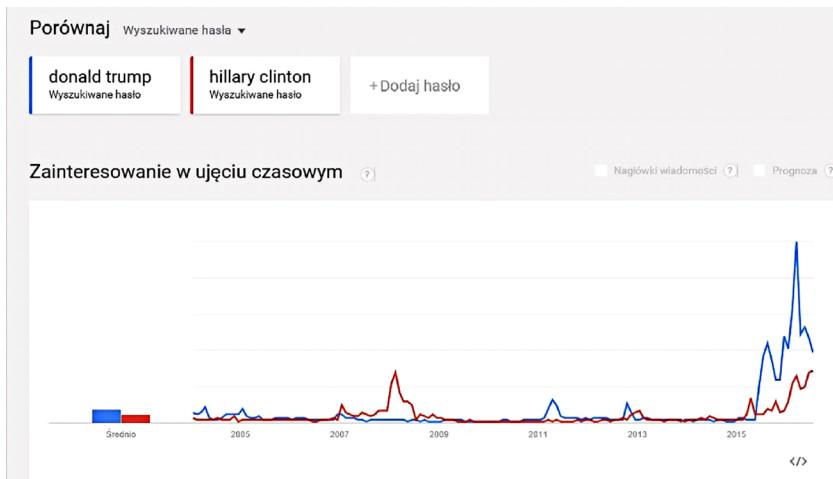
quences are widely discussed in *The Handbook of the Psychology of Communication Technology* (Sundar 2015). The media's audience, and the ways in which it "consumes" media products, have also changed. An international team of authors has analysed the dynamics of these processes in *The Handbook of Media Audiences* (Nightingale et al. 2011). The reader may learn more about the digitalisation of media, the creation of network content by users themselves and the need to change the measurements used in market and media research in the *Handbook of Research on Digital Media and Advertising: User Generated Content Consumption* (Eastin et al. 2011). Social relationships, especially in the so-called social networks, have also been transformed. The Internet 2.0 has created increasing opportunities to research social dynamics and structures supported by Web networks. This has become possible for many reasons, including intelligent research and programming tools. The authors of *Computational Social Network Analysis* (Abraham, Hassaniem and Snášel 2010) describe the design of research and functions of selected applications. R. Poynter further reviews online research and applied techniques. The researcher presents opportunities for carrying out research that is quantitative (surveys) and qualitative (among other projects, online research focus groups). He also explains the details of blog and social media research (Poynter 2010). Their role has been constantly growing, and today almost no one questions it. Scientists have been attempting to analyse and explain the influence of social media on modern communities, and to identify new research methods through the use of innovative tools (Rogers 2013; Weller, Axel et al. 2014). One of the publications in this field is *Digital Methods for Social Science*. This work, as its editors write, uncovers innovations in the area of digital social research. They emphasise that digital methods present a valuable and increasingly integral part of a social researcher's toolkit (Snee et al. 2016, 1). For a description of new technologies, an analysis of data and computer research concentrated on the impact of the media on the public, the reader may turn to *Human-Centered Social Media Analytics* (Fu et al. 2014). The *Data Journalism Handbook* (Gray, Bounggru, Chambers 2012) deals with sources of information for journalists, and forms of data processing. Those interested in a description of the subject should also acquaint themselves with related academic journals such as: *Communication Research*, *Journal of Computer-Mediated Communica-*

tion, Information, Communication & Society, Global Media and Communication, Digital Journalism, Big Data & Society and Social Media + Society.

Examples of tools and their use

One of the simplest free-access tools is *Google Trends* (<https://www.google.pl/trends/explore>). It works as a browser, not demanding a log-in or any technical knowledge. On the other hand, it provides quick access to content that interests Internet users, in terms of both time and region. After researched content is entered in the search engine (concept/password/name), the obtained results are presented on a timeline (Chart 2) and a map of the countries where the subject is most often queried. Within a few seconds, it is possible to check how much interest international issues raise in various parts of the world. Additionally, depending on the subject of the research, news headlines (articles) are also displayed, and so Internet users are as engaged as possible in the subject. *Google Trends* also collates the context of the query. This very important capability makes it possible for qualitative conclusions to be drawn from quantitative data. For example, in 2016 queries about the U.S. Republican candidate had a different context than those about his rival. The name “D. Trump” was most often combined with the items “president”, “worth” and “wife”. H. Clinton was searched on the net in combination with queries about other politicians, such as “Trump”, “Sanders” and “Bill Clinton”. “Clinton – president” was not greatly popular. Such data hold a mirror to the mood of a society, and the real interests of Internet users.

Chart 2. Interest in Donald Trump and Hillary Clinton, over time



Source: <https://www.google.pl/trends/explore#q=donald%20trump%2C%20hillary%20clinton&cmpt=q&tz=Etc%2FGMT-2>; accessed 10.07.2016.

A similar possibility is offered by *Tweetreach.com* with the difference that the tool serves only as an analysis of Twitter. In recent years, Twitter has become a very popular communication platform. It has awoken significant interest among not only Internet users, but also traditional media. Thanks to the functionality of the service, it can make it easier for journalists and media scholars to search for the current subjects of discussions among users, and to follow them and the subjects they raise. Twitter, by adopting the hashtag content standard, has allowed for the rapid searching, sorting and expanding of the scope of information, even as data is gathered. This is essential, because important events and incidents are marked in the same way, regardless of the place of publication or the nationality of the authors. A further advantage is presented by the Twitter API procedures³, which make Twitter more open and accessible than other social platforms. This, in turn, makes it easier for programmers to develop tools for data analysis, increasing the accessibility of tools for researchers, who do not always need to have technical knowledge (Ahmed 2015). Examples include *Trends24.in*, *Tweetreach*.

³ These may be understood as a precisely defined set of regulations and descriptions by which computer programs communicate with one another. Data collection takes place according to what is required and permitted by a platform/portal (in this case, Twitter).

com and *KeyHole.co.*, all of which work from a browser. *Trends24.in* does not even require a Twitter account. The application helps in following trends (most often using hashtags and key words) on a global scale or within a country (or a given city). Results are presented graphically, each hour, around the clock.

Most of the available data is created according to the needs of the market (advertising, public relations and so on). These applications generate a range of data, often necessary from the point of view of academic research or difficult to understand for novices in a given field. In response, the following tools have been created by an international academic group (Table 1).

Table 1. List of academic Internet research tools for social scientists

Researched Tool	System	Source	Platform researched
Mozdeh	Windows	http://mozdeh.wlv.ac.uk/installation.html	Twitter
Webometric Analyst	Windows	http://lexiurl.wlv.ac.uk/	Twitter (with image extraction capabilities) YouTube Flickr Mendeley Other Internet resources
NodeXL	Windows	http://nodexl.codeplex.com/	Twitter YouTube Flickr
Netlytic	web-based	https://netlytic.org/	Twitter, Facebook, YouTube, Instagram
Twitter Arching Google Spreadsheet (TAGS)	web-based	https://tags.hawksey.info/	Twitter

Researched Tool	System	Source	Platform researched
Chorus	Windows	http://chorus-analytics.co.uk/chorus/request_download.php	Twitter
COSMOS Project	MAC OS X, Windows	http://www.cosmosproject.net/ http://socialdatalab.net/software	Twitter
Visibrain Focus	web-based	http://www.visibrain.com/en/	Twitter
DiscoverText (30-day trial, free of charge)	web-based	http://discovertext.com/	Twitter, Facebook, Blogs, Forums, Online news platforms

Source: Ahmed, 2015.

The tools listed above do not require their users to have technical knowledge. Often, however, it is necessary to use data analysis software, such as Excel, SPSS and Statistica, to make full use of their potential. A basic knowledge of statistics can further facilitate the interpretation of obtained results.

Mozdeh, a product of the Statistical Cybermetrics Research Group at Wolverhampton University, arose as part of two European Union projects, CREEN and CyberEmotions. It serves the collection, archiving and complex analysis of Twitter entries. After the program is installed on a computer disk, the user may carry out quantitative and qualitative analysis: tweet counts, time series analysis, word frequency counts, the creation of graphs and calculations for social networks, sentiment analysis and other operations (Chart 3). An essential function is the choice of language, facilitating the analysis of the international aspect of a given subject.

Chart 3: Mozdeh Application Window

The screenshot shows the Mozdeh application interface. At the top, there's a menu bar with File, Analyse, Networks, Data, Subprojects, and Info. Below the menu is a toolbar with icons for big data, Boolean Search (or), Next, Sort by: Date (as), Show original, Advanced Search, Save, Spam Filtering, and Co-word comparisons.

Advanced Search: Set to 'big data'. It includes fields for ID of first item to show (1), First date to show (2016.7.20), Results to show (38), and checkboxes for Hide hashtags and @s, Hide duplicates (v. slow), and Only return results from this topic/label tweeters (case sensitive).

Advanced Search Tips: Describes how AND, OR, and NOT operators work, along with examples for using hashtags and quotes.

Results only from this Tweeter/feed ID: A dropdown menu. Below it are checkboxes for Only tweeted by label tweeter, Select alternative subproject, and a sentiment analysis section with '<=> +ve sentiment <=>' and AND/OR buttons.

Filtering: Options include Any gender, This text must be in Twitter usernames, Exclude matching tweeters instead, and Exclude tweeters with this many tweets collected (0).

Results: A table with columns for ItemID, Date, and Text. The table contains 15 rows of tweet data, such as '000001[1] | 2016.7.20 | cloud native big data partner share their experience' and '000015[1] | 2016.7.20 | rt @analyticsfr way retailer are using big data and hadoop #bigdata #analytic'.

Feedback: On the right side, there's a feedback box with 'Feed #19 Info', 'URL/Source', 'Topic/tweets', 'Author: Latent', 'Retw.', 'L=laststart 30', 'Cles 0--> Big', and 'know you better'.

Source: the author's work.

A further application useful in sentiment analysis is SentiStrength, not listed in the table. It is available free of charge if it is used for academic purposes. “Opinion-mining” research aims to verify positive or negative tones in unstructured texts (Pang, Lee 2008). SentiStrength makes it possible to analyse texts in different languages, including – and this is a rarity – those in Polish. The free version works in a Windows environment. The reader may find more information on the use and possibilities of the application in reviewed journals (Thelwall et al. 2012, 163-173), and on the project authors’ web page.

Webometric Analyst is a fairly extensive Internet tool created, like *Mozdeh*, by the *Statistical Cybermetrics Research Group*. It helps make possible the analysis not only of Twitter, but also of other social media (YouTube, Flickr). A characteristic feature is that it researches the impact of documents, Internet pages and specialised search engines such as *Google Books* and *Mendeley*. It is a very useful tool for webometric measurements: quantitative research of Internet pages and their elements, and of the words that appear there, hyperlinks and Internet search engine results (Thelwall 2009, 1). Hyperlinks are researched for the purpose of getting to know their types, significance and impact in an

electronic information environment (Skalska 2014, 160-161). Thanks to specialist online measurements, precise research becomes possible of what interests Internet users, which content gains their attention and which elements they share.

NodeXL was created by *The Social Media Research Foundation*, whose mission is to develop software for social media analysis. *NodeXL* serves to add or extend MS Excel. When the program is started, it appears as a new tab with extended functionality. The basic version is available free of charge, but contains only elementary steps. The PRO version offers advanced functions for the importing of data from social media, calculations involving social network indicators, tone analysis and the publication of reports. It also enables visualisation of the social networks that are studied.

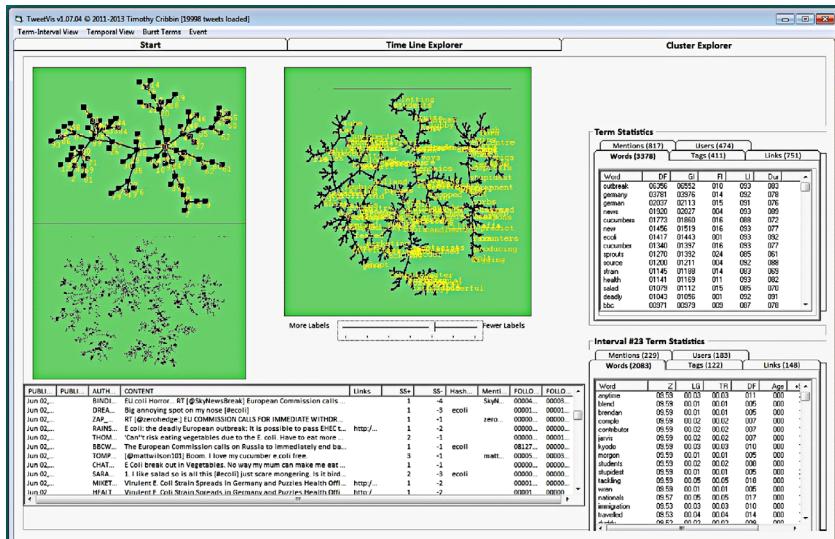
A further tool for rapid analysis of Twitter, *TAGS*, was created privately by M. Hawksey. It is available in the form of a Google Sheets template, free of charge. The application makes possible the automated configuration and collection of search results.

Netlytic is an extensive online text and social network analyser, used to collect (download) data from various social media (Twitter, Facebook, YouTube, Instagram, RSS feed readers), as well as texts and csv files. The application makes it possible to verify what topics are popular, find conversations and analyse their threads. In addition, online social networks may be built, visualised and analysed. A particularly interesting capability is that of geolocated data (geo-coded social media data). The program creators have promoted their product as a tool made by researchers and for researchers, although a user does not require computer knowledge. Twitter may also be researched using geolocation with another free-of-charge academic tool, *iScience Maps*. It works in a browser, not requiring the installation of additional elements. It has two built-in modules, one of which allows global searches and long-term trend analysis; the other is intended for local searches and comparative research of specific concepts and places over a given period of time. The discussion of a given event from a particular distance (in kilometres) may be observed; for example, it is possible to analyse entries after a terrorist attack at the place where the tragedy occurred, and within a radius of 50 kilometres (Reips, Garaizar 2011).

Chorus and *COSMOS* are two further, well-developed academic

tools. Both serve to deepen analyses of Twitter content. A user may gain access, free of charge, by stating their wish to use the service, and the scope of their research. The platform was created to allow Twitter to be used to develop social research. It has made it easier to manage projects and analyse information in the area of Big Data (Chart 4). P. Brooker, in the journal *Big Data & Society* (2016) has written extensively on ways to carry out such research and the need to visualise data. COSMOS was developed between 2012 and 2015 by *The Collaborative Online Social Media Observatory*, a consortium created by Cardiff University, the University of Warwick and the University of St. Andrews. Representatives of the social and technical sciences, such as statisticians, mathematicians and health sciences experts, developed the initiative. Its main task has been to develop the methodological, theoretical, empirical and technical foundations of research into social media. The project now continues as *The Social Data Science Lab* at Cardiff University. The work environment is OS and Linux.

Chart 4. Chorus Interface



Source: <https://digitalmethodsnmi.files.wordpress.com/2013/04/pic2.jpg>.

Visibran, listed in the table, is a commercial tool that allows social media to be monitored in real time. *Brand24* deserves mention as its Polish

counterpart, with an international reach and similar functionality. This tool works in a Polish-language version (*Brand24.pl*). *Brand24.com*, in turn, allows monitoring on a global scale. The tool collects data from numerous social media, as well as Internet pages and forums, then automatically carries out sentiment analysis of posts. It presents changes in the intensity of numerous mentions, and their range. Properly set filtering allows for the monitoring of narrowly specified data. *Brand24* also creates PDF reports automatically, exports data to .xls files and generates information graphs of the most important statistics. An indicator of author influence scores has also been developed, classifying the source of results. The indicator value is correlated with the range, popularity, activity and interactivity of the researched source. *Brand24* provides this tool on a commercial basis, but the company works with various academic centres where the tool is used for educational purposes. Further, everyone may access the service free of charge for two weeks. In the Polish literature on the subject, the first academic research that appeared used the *Brand24* tool. Examples of such research are described in publications by A. Paluch and B. Łódzki (2016), and Łódzki (2016, 2017).

At this point, other Polish commercial analytical tools, worth using in research, deserve mention. SentiOneUnamo (a joint effort of the companies Postonly and Monitori) and Sotrender offer their tools free of charge for a trial period, usually two weeks. With SentiOne, data may be collected on any topic that appears on the Internet. Social media, microblogs, Internet portals, forums and blogs are monitored. The tool counts and analyses Internet posts. A further possibility is the automatic production of research reports. Unamo, on the other hand, besides monitoring Internet mentions, is also a specialised tool for monitoring the ranking of phrases or keywords that users select in looking for a given service. Sotrender, on the other hand, compares ways in which profiles are managed (how they are edited), and analyses the behaviour of receivers (users). Further, the tool generates reports and automatic recommendations (of what may be improved in the analysed profiles).

Texifter serves research concentrating on text analysis. This application, created by Dr. S. Stulman, is offered on a commercial basis (with one-month access free of charge). Researchers and students may receive discounts of between 50 and 75%. An alternative may be *Voyant*, a free-of-charge academic tool. Working as a browser, it can analyse both in-

dividual texts and whole groups of articles. Simply and quickly, it can analyse the most frequently occurring words, mapping them and verifying their context (the sentences in which they appear). The tool offers automatic visualisation of results and their export to external applications (<https://voyant-tools.org/docs/#!/guide/start>).

One of the few academic tools for Facebook research is Netvizz. This application can be installed on a website user's private account. It allows a researcher to extract data using an API protocol. The software allows operations in several areas. One is the analysis of open groups. The tool creates networks of user connections, and tabular files of activity around posts published in groups. The same operations are possible on web pages (open profiles). It is also possible to analyse networks of connections between web pages. The tool also has an internal search engine. The interface allows pages and groups, devoted to particular issues, to be identified and examined. Netvizz automatically performs a statistical count of the links that users share. The tabular data contain a range of information that allows the observation of the activity of search profiles and their users (readers). The tool identifies types of published content (whether an item is a post, photo, video or link). It counts the number of "likes", shares and comments. These features constitute the portal's most important use (Rieder 2013, 346-355). In 2016, Facebook introduced supplementary icons for expressing emotions and users' relationship to published content. The current version can also analyse these elements. A further available option analyses the 200 best posts, or only pictures. The "Timeline Images" model creates a list of all pictures from the "Timeline Photos" album on the site, and counts all reactions. Netvizz also provides an internal search engine which can help profiles containing indicated keywords to be found. Furthermore, the software can produce data for further processing, not only in spreadsheets but also programming for network analysis, such as Gephi.

Sources of information on search tools

A short overview of selected academic and commercial tools does not exhaust the subject. Those who are interested can deepen their knowl-

edge using numerous Internet sources. One of the most interesting is the website of the Digital Methods Initiative, a research group (<https://wiki.digitalmethods.net/>), which contains a rich compendium of analytical knowledge. The group organises summer schools, during which users are introduced to the secrets of the latest research. The project is carried out at the University of Amsterdam.

On the site <http://socialmediadata.wikidot.com/>, D. Freelon collects information about analytical tools. The researcher's list of applications and programs has been divided according to the level of skill required to use them. The list also contains applications that do not require specialist knowledge.

The Internet search engine of the <http://dirtdirectory.org/> application is also very useful. The scope of research is chosen on the main catalogue page, and more detailed search criteria are displayed: work environment, costs, types of licensing and more precise subjects. The results are presented in English and Spanish.

Chart 5. DIRT – Internet search engine of research tools

The screenshot shows the homepage of the DIRT (Digital Research Tools) website. At the top, there is a navigation bar with links for 'About', 'Tools', 'Contribute', and 'Users'. Below the navigation bar is the DIRT logo, which features a stylized lightbulb icon with a plant growing out of it, followed by the text 'DIRT' in large, bold, black letters, and 'Digital Research Tools' in smaller text below it. To the right of the logo is a search bar with a placeholder 'Search' and a magnifying glass icon. On the left side of the main content area, there is a section titled 'Analysis' with several dropdown menus for filtering search results: 'Platform', 'Cost', 'Exclude', and 'License'. Below these filters is a 'Research objects' dropdown menu. To the right of the filters are buttons for 'Sort by' (set to 'Updated') and 'Order' (set to 'Descending'), along with a 'Reset' button. A question 'What kind of data should the tool work with?' is also present. On the right side of the page, there are two sections: 'LANGUAGES' (with options for English and Español) and 'RESEARCH ACTIVITIES' (with a dropdown menu currently showing 'Capture (74)').

Source: <http://dirtdirectory.org/>.

It is also possible to use the application for managing Twitter content, such as TweetDeck, and with its help to search for interesting research of the content (#bigdata, #researchtool).

Conclusions

The presented tools and Internet sources may guide those who are beginning their adventure in media studies research in the digital realm. They do not entirely replace traditional methods and tools. They do, however, allow the research field to be widened, and projects to be carried out with a wider scope. The author encourages readers to get to know and to test the applications, which are increasingly being used in numerous academic centres. Those who know and use them in a practical way are better able to raise the ranking of their publications, and to analyse and compare their work with other research. Through the use of these tools, a great deal of trans-national research has begun, such as an analysis of the ways in which citizens communicated during an election campaign in Nigeria (Bartlett, Krasodomski-Jones 2015), research into the problem of international cybercrime (Williams, Wall 2013), ways in which sports events are reported (Tour de France) and the reaction of Twitter users to natural catastrophes such as floods and earthquakes (Weler, Bruns 2014). G. Barnett and his team (2017) analysed public opinion related to relations between nations, as revealed by conversations in social media: over 1.8 billion Facebook posts in English, and 1.51 billion in Chinese on Weibo, were used in the study. The research potential is almost unlimited. In the era of Big Data, with ready access to ever-richer databases and information, scholars in the humanities should be open to new, interesting research fields. It is, in a sense, an irreversible process. As F. D'Orazio notes (2013), social data in the era of social media has become “qualitative data on a quantitative scale”.

References:

- Abraham A., Hassaniem A., Snášel V. (2010), *Computational Social Network Analysis. Trends, Tools and Research Advances*, London: Springer.
- Ahmed W. (2015), *Using Twitter as a Data Source: An Overview of Current Social Media Research Tools*, <http://blogs.lse.ac.uk/impactofsocialsciences/2015/07/10/social-media-research-tools-overview>, accessed 10.07.2016.
- Armerding T. (2016), *Big Data and Elections: The Candidates Know You – Better Than You Know Them*, <http://www.cssoonline.com/article/3095709/data-protection/big-data-and-elections-the-candidates-know-you-better-than-you-know-them.html>, accessed 14.07.2016.

- Barnett G.A., Xu W.W., Chu J., Jiang K., Huh C., Park J.Y., Park H.W. (2017), *Measuring International Relations in Social Media Conversations*, "Government Information Quarterly1" (34), pp. 37-44.
- Bartlett J., Daniel N., Fisher A., Jesperson S., Krasodomski-Jones A. (2015), *Social Media for Election Communication and Monitoring in Nigeria*, London: Demos.
- Brooker P., Barnett J., Cribbin T. (2016), *Doing Social Media Analytics*, "Big Data & Society 3" (2), pp. 1-12.
- Corner J. (ed.) (2005), *International Media Research. A Critical Survey*, New York: Routledge.
- D'Orazio F. (2013), *The Future of Social Media Research*, <http://abc3d.tumblr.com/post/62887759854/social-data-intelligence>, accessed 16.05.2016.
- Eastin M. Daugherty, T. Eastin, Matthew S. (2011), *Handbook of Research on Digital Media and Advertising: User Generated Content Consumption*, Hershey, New York: Information Science Reference.
- Elliott T. (2013), *7 Definitions of Big Data You Should Know About*, <http://timoelliott.com/blog/2013/07/7-definitions-of-big-data-you-should-know-about.html>, accessed 17.05.2016.
- Farrell D., Petersen J. (2010), *The Growth of Internet Research Methods and the Reluctant Sociologist*, "Sociological Inquiry", 80 (1), pp. 114-125.
- Foote K. (2017), *A Brief History of Big Data*, <http://www.dataversity.net/brief-history-big-data>, accessed 20.12.2017.
- Fu Y. (ed.) (2014), *Human-Centered Social Media Analytics*, Switzerland: Springer.
- Goggin G., McLelland M. (eds.) (2009) , *Internationalizing Internet Studies*, New York: Routledge.
- Gray J., Bounegru L. & Chambers L. (2012), *The Data Journalism Handbook*, Cambridge: O'Reilly.
- Halevi G., Moed H. (2012), *The Evolution of Big Data as a Research and Scientific Topic: Overview of the Literature*, "Research Trends", 09, 3-6, https://www.researchtrends.com/wpcontent/uploads/2012/09/Research_Trends_Issue30.pdf, accessed 15.05.2016.
- Lódzki B. (2016), *Aktywność mediów tradycyjnych na platformach społecznościowych podczas kampanii do Europarlamentu w 2014 roku – czynny moderator dyskusji politycznych czy skuteczny autopromotor?*, in: Furman W., Kuca P., Szczepański D. (eds.), *Media w kampaniach wyborczych*, Rzeszów: Wyd. Uniwersytetu Rzeszowskiego, pp. 171-181.
- Lódzki B., Paluch A. (2016), *Analysis of Intermedia Agenda-Setting: the Case of the 2014 EP Election in Poland*, In Boicu R., Branea S., Stefanel A. (eds.), *Political Communication and European Parliamentary Elections in Times of Crisis*, London: Palgrave Macmillan, pp. 43-56.

- Łódzki B. (2017), *Analiza aktywności aktorów politycznych i mediów tradycyjnych w mediach społecznościowych w trakcie kampanii prezydenckiej 2015 w Polsce – studium przypadku*, in: Kolczyński M. (ed.), *Polskie wybory 2014-2015. Kontekst krajowy i międzynarodowy – przebieg rywalizacji – konsekwencje polityczne*, Katowice: Wydawnictwo Uniwersytetu Śląskiego, pp. 168-185.
- Nielsen (2012), *State of the Media. The Social Media Report*, <http://www.nielsen.com/us/en/insights/reports/2012/state-of-the-media-the-social-media-report-2012.html>, accessed 31.03.2016.
- Nightingale V. (ed.) (2011), *The Handbook of Media Audiences*, Oxford: Wiley-Blackwell.
- Pang B., Lee L. (2008), *Opinion Mining and Sentiment Analysis*, “Foundations and Trends in Information Retrieval”, 1(1-2), 1-135.
- Poynter R. (2010), *The Handbook of Online and Social Media Research: Tools and Techniques for Market Researchers*, Wiley.
- Reips U.-D., Garaizar P. (2011), *Mining Twitter: Microblogging as a Source for Psychological Wisdom of the Crowds*, “Behavior Research Methods”, 43, 635-642. doi <http://dx.doi.org/10.3758/s13428-011-0116-6>.
- Rieder B. (2013), *Studying Facebook via Data Extraction: The Netvizz Application*, in: WebSci ‘13 Proceedings of the 5th Annual ACM Web Science Conference, ACM, New York.
- Rogers R. (2013), *Digital Methods*, London: The MIT Press.
- Shearer E., Gottfried J. (2017), *News Use Across Social Media Platforms 2017*, <http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017>, accessed 20.01.2018.
- Skalska-Zlat M. (2004), *Cybermetrics, Netometrics, Webometrics – nowe pojęcia i zadania informetrii*, in: Kocójowa M. (ed.), *Przestrzeń informacji i komunikacji społecznej*, Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, pp. 159-168.
- Roberts S., Snee H., Hine C., Morey Y., Watson H. (eds.) (2016), *Digital Methods for Social Science. An Interdisciplinary Guide to Research Innovation*, London: Palgrave.
- Sundar S. (ed.) (2015), *The Handbook of the Psychology of Communication Technology*, Oxford: Wiley.
- Szews P. (2014), *Serwisy społecznościowe jako źródło informacji dziennikarskiej – problemy i zagrożenia*, “Kultura, Media, Teologia”, (18), pp. 90-104.
- Thelwall M. (2009), *Introduction to Webometrics: Quantitative Web Research for the Social Sciences. Synthesis Lectures on Information Concepts, Retrieval, and Services*, Chapel Hill: Morgan & Claypool Publishers.
- Thelwall M., Buckley K., Paltoglou G. (2012), *Sentiment Strength Detection For The Social Web*, “Journal of the American Society for Information Science and Technology”, 63(1), pp. 163-173.

Wall D., Williams M. (2013), *Policing Cybercrime: Networked and Social Media Technologies and the Challenges for Policing*, “Policing and Society”, 63 (4), pp. 409-412.

Weller K., Bruns A., Burgess J., Mahrt M., Puschmann C. (2014), *Twitter and Society*, New York: Peter Lang.

Darknet and Medical Big Data. Deep Internet as a Space for Illegal Trade in Medical Information

Abstract: The article discusses issues that affect phenomena such as Big Data and Darknet space, which overlap in the aspect of theft and illicit trade in sensitive medical data. An attempt will be made to answer questions about the boundaries of freedom and anonymity of the global Internet network and the security of large data collections collected and analyzed by institutions in the health care sector. The starting point will be the assumption that medical documentation is the most comprehensive collection of information about the identity of a person, which can now be found in the resources of the Internet. The study will analyse the properties of Darknet, the path of Internet criminals will be tracked, and the analysis of websites and links leading to Darknet's trading websites will show the simplicity of the entire procedure and the potential profits brought by these activities.

Keywords: Big Data, Darknet, medical data, Digital Humanities, sensitive data.

Introduction

Recently, the concept of the so-called Big Data has become more and more popular, which in the simplest way can be defined as those that will be characterized by volume, diversity and variability. Their analysis goes far beyond simple operations on one given type of information and includes aspects of data analysis, such as the generation of hypotheses, and not simply testing them. Willy-nilly, the era of great data has come, and computer scientists, physicists, economists, mathematicians,

political scientists, bioinformatics, sociologists and other scientists are demanding access to a huge amount of information generated by people, things and their mutual interactions. There will be arguments about the potential benefits and costs of genetic sequence analysis, interactions in social media, health data, data produced by states and governments, and other digital traces left by people (Boyd, Crawford 2012).

Parallel to the large collection of data gathered by various institutions, we are dealing with the development of the practice, which may call into question the legitimacy of storing such diverse and sensitive information. The phenomenon of massive theft of huge amounts of data is becoming a global plague. Beginning with hacking into banking systems, hackers set themselves the goal of stealing personal information, data that is secret, because they concern the most personal issues of our existence. The problem of continuously implemented security systems is not the main theme of this study, but undoubtedly this issue leaves room for further analysis. At the meeting point of Big Data and the procedure of stealing them, there is a space for the exchange of these data, questions arise as to why medical data of patients are such an interesting goal and what are the interests of those who want to get into them. How is this data exchanged and for whom can they be interesting? In order to be able to answer these questions, it is necessary to scratch the trade space of the said data, and the place that perfectly meets the criminal requirements, constituting an excellent outlet, is the hidden network (Darknet). Firstly, one should look at the characteristics of Big Data, determine what they are, why they are collected and why their analysis is so important from the point of view of improving medical services and public health standards. Next, it is worth discussing the anonymous space in which free trade of these data takes place. Finally, it is also worth trying to give an answer to the question, who gets the most benefit from trading data in the hidden Internet space and why this procedure is so lucrative. For this purpose, the results of previous research in this area will be used (Van Buskirk 2016, Sutton 2013, Pang 2016, Martin 2014 Bancroft 2016, Wehinger 2011), as well as sources and press reports and analysis of websites on which one can find references and links leading directly to the Silk Road, which is a place of data trade in the area of the Darknet.

Big Data. General characteristics and definitional and methodological problems

As claimed by Danah Boyd and Kate Crawford in their article *Critical Questions for Big Data. Provocations for a Cultural, Technological, and Scholarly Phenomenon*, Big Data is a weak term for at least several reasons. Citing even the assumption of Lew Manovich (Gold 2011), the authors indicate that this term was originally used in science to refer to data sets large enough to require the use of supercomputers. Nevertheless, what once required such machines can now be analyzed on a personal computer with standard software. There is no doubt that the amount of data currently available for analysis is huge, but this is not a distinctive feature for a specific, completely new ecosystem of data. The Big Data analysis focuses on the temporal stability of their interrelationship, not on the search for cause-and-effect relationships, and the assumptions about the probability distribution often on the occasion of analysis are not required (Lee, Yoon 2017).

Given the above-mentioned matter, it is impossible to indicate a definition that would not contain interdisciplinary aspects and would not describe Big Data as a technological, cultural phenomenon, and at the same time subject to scientific analysis. Therefore, firstly we have technology that allows to maximize computing power of computers (being analytical tools) and algorithmic accuracy that allows analyzing, combining and comparing large sets of collected data. In addition, we deal with in-depth, scientific analysis carried out on large data sets to identify patterns and better quality models of economic, social, technical and legal phenomena. Thirdly, we are dealing with the creation of something in the form of a cultural, universal belief that large data sets offer a higher form of intelligence and knowledge that can help generate observations and conclusions that would be impossible before. All this is based on the aura of truth, objectivity and accuracy (Boyd, Crawford 2012).

Similarly as other socio-technical phenomena, Big Data triggers both utopian and dystopian rhetoric. On the one hand, Big Data is seen as a powerful tool for solving various social problems, offering insights into areas as diverse as research on cancer, terrorism and climate change. On the other hand, they are presented in terms of the disturbing manifes-

tation of the Big Brother, enabling invasion into the sphere of privacy, reduction of civil liberties and increased state and corporate control. As with all socio-technical phenomena, overlapping visions of hope and fear often obscure the more nuanced and subtle changes that are just taking place (Boyd, Crawford 2012).

The manner in which we deal with the emergence of the Big Data era is crucial, because this phenomenon occurs in the environment of often extreme uncertainty and rapid change, and currently decisions made in this respect will shape the future to a large extent. Due to the increased automation of data collection and analysis – as well as algorithms that can be distinguished and illustrated in terms of large-scale human behavioral patterns – one should ask which systems direct and regulate these practices. Lawrence Lessig in his book *Code and Other Laws of Cyberspace* distinguished four factors regulating the behavior of people on the Internet: code, social norms, law and market (Lessig 2006). All these four regulators are interlinked, they are interdependent, together they create the conditions that make up the environment in which Internet users operate. All of them are relatively independent of the individual, they are something that a person finds as data when they want to communicate with others through this worldwide network of connected computers. But at the same time they are created by people, and their shape is not something imposed in advance, immanently inherent in the nature of the world. Two of the above factors can be intentionally changed and transformed: it is a code (in other words the architecture of the Internet) and the law. By means of the changes of the two above elements, we can influence the shape of the market and social norms. But at the same time the market – understood as a group of commercial entities, such as film, music or other market players – exerts pressure on politicians to change the law in such a way as to help these entities achieve the greatest financial benefit (FDC 2017) When we are dealing with Big Data, however, these four forces are often contradictory. The market “sees” Big Data as a chance for quick profit, namely marketers use it to target advertising, suppliers use it to optimize their offer, and Wall Street bankers use it for market analysis. Legislation has already been proposed in many countries to limit the collection, storage and processing of data, usually in connection with privacy concerns (an example can be here the US Do Not Track Online Act of 2011) (Wolniak

2006). Features such as personalization allow quick access to more relevant information, but also give rise to difficult ethical questions and raise social anxiety.

It is clearly visible that the issue of Big Data is not a homogeneous area. As we have already mentioned, this is a highly diverse and interdisciplinary field. Extracting and analyzing data that can be used to create models of a certain slice of reality is not a simple task. Big Data offers a new way for humanities to achieve a higher status of science, as the quantitative method and the emphasized objectivity of research results are successful. This makes many social spaces, thanks to the possibility of analyzing large data sets, measurable spaces. Nevertheless, this enthusiasm should be somewhat dimmed. In fact, working with Big Data is still a subjective action, and what is quantified does not necessarily have to be closer to the objective truth – especially in the case of analysed messages sent from social networking sites. However, the mistaken belief is that researchers using qualitative methods deal with the interpretation of stories, and researchers using quantitative methods, so-called hard data, deal with researching facts and building strong theories. In this way, Big Data and related research are inherently embedded in long-established divisions, which only deepened during long academic debates about the scientific method and the legitimacy of social sciences and humanistic investigations.

Debates on the concept of objectivity in science have been taking place in principle since the beginning of social sciences, and at the end of the nineteenth century it was one of the central issues for the philosophy of science. Objectivity theorems suggest belonging to the sphere of objects, to things that exist in themselves and for themselves. On the other hand, subjectivity is perceived with suspicion, as it is in case of other forms of individual and social conditioning. The scientific method tries to get out of the subjective domain by applying a dispassionate process in which the hypotheses are proposed and tested, which ultimately is supposed to lead to the improvement of knowledge.

Nevertheless, claims about objectivity are again made by subjects and are based on subjective observations and choices, because all researchers are largely data interpreters. They decide which section of reality they will analyze, what sample they will analyze and what data they will treat as more relevant to the subject of their research. As indicated by Barbara

Krauz-Mozer controversial discussions about these assumptions are not a dispute about a research strategy, but rather a kind of “social game” in which the parties can submit their arguments for a long time without gaining any final approval or objection for them, because they participate in a dispute regarding issues that are essentially undecidable. At most, joining this game will divide the participating researchers into proponents of naturalism or anti-naturism (methodological dispute), followers of objectivity or constructivism (ontological dispute), for those who consider valuing as being inalienable in science and those who advocate its neutrality (epistemological dispute) (...) (2010). This is an eternal problem of valuation in science, which cannot be effectively escaped even using methods based on the analysis of powerful data sets, such as Big Data. As Lisa Gitelman notes, data first need to be imagined as data, and this process of imagining data is connected with interpretive basis, because each discipline has its own norms and standards for imagining data (Gitelman 2011).

Scientists dealing with information technology have begun to engage in the activities of social sciences, so there is a tendency to claim that their work comes down to the analysis of facts, not their interpretation. The model can be mathematically correct, the experiment may seem right, but as soon as the researcher tries to understand what it means, the process of interpretation has begun. This does not mean that all interpretations are equal, but rather that not all numbers should be treated as neutral. The very decision of the researcher, already at the initial stage of designing the research task, defining what will be measured, also results from its interpretation. For example, in the case of data from social media, there is a process of so-called “data cleansing”. It is already deciding which attributes and variables will be counted and which will be ignored. This process is inherently subjective (Boyd, Crawford 2012). Therefore, despite the importance of Big Data for the development of science, their role should not be overestimated and the contribution of their analysis should not be absolutized.

Big Data – medical data

The recent and rapid increase in the generation of digital data and the equally rapid development of advanced computing science allow us to

draw new conclusions from huge data sets in various fields. In the field of health care, discovering new, useful to the action insights was not so common, although several success stories were published in the media and academic journals. This delayed advancement of Big Data sets technology in the healthcare sector is somewhat strange given the earlier forecast that the use of this technology is inevitable and that the health sector will be the one that is expected to benefit most from their analysis (Hutchings 2017). The growing gap between health care costs and expenses is one of the most important issues, and many efforts to fill this gap are ongoing, especially in developed countries. The gap between health care costs and revenues was identified as the result of poor management of research conclusions, poor use of available analyses, which in turn led to loss of opportunities, wasted resources and generating potential damage to patients. It was suggested that this gap could be overcome by developing a “continuous health care system” in which a close connection between research and operational health care is created, so that the collected data could be used effectively.

Medical Big Data, as the material subjected to analysis, have different characteristics that differ not only from Big Data from other disciplines, but also differ from traditional clinical epidemiology. Big Data technology has many areas of application in healthcare, including such as predictive modeling and clinical decision support, disease and safety monitoring, public health and more. Big Data analytics often uses methods developed in data mining, including classification, cluster creation and regression (Rosa, Chen, Binder 2015).

The complexity of health care systems is largely due to the diversity of health problems and concomitant diseases, the heterogeneity of treatment and the results, and the complexity of research projects, analytical methods and approaches to collecting, processing and interpreting data on healthcare. There are various sources of big medical data, such as registration of administrative claims, clinical records, electronic health cards, biometrics, patient data, Internet, medical imaging, biomarker data, prospective cohort studies and large clinical trials (Slobogean, Giannoudis, Frihagen, Forte, Morshed, Bhandari 2015).

Big medical data has several characteristic features that differ from Big Data from other areas. They are collected on the basis of protocols (i.e. permanent forms) and are relatively ordered, partly due to the extraction

process, which simplifies the processing of primary data. Another important feature is that medicine is practiced in a critical security context in which decision-making should be undertaken with extreme caution. Big medical data, their organization and analysis can be costly activities due to the additional involvement of personnel, the use of expensive equipment and the potential discomfort of the patients involved (Lee, Yoon 2017).

The potential value of big medical data has been demonstrated in: 1) provision of personalized medical services; 2) the use of clinical decision support systems, such as automated medical image analysis and searching for relevant medical literature; 3) adjusting diagnostic and treatment decisions as well as educational messages in order to support desired patient behavior with the help of mobile devices; 4) analysis of population health based on Big Data, revealing patterns that may have been omitted if smaller batches of uniformly formatted data were analyzed instead; and 5) detection and prevention of fraud. Diagnosis based on high resolution measurements, such as microarray or next-generation sequencing, monitoring of molecular characteristics during treatment and continuous monitoring of health are potential applications of big medical data. (Lee, Yoon 2017)

Medical data and their analysis are therefore an extremely valuable source of information, which, however, is not always used for the purposes described above. Increasingly, we are dealing with a situation of “outflow” of huge data sets from servers of medical units on which they are collected and stored. Due to the growing value of sensitive data, theft of this information by hackers has become a popular practice. But what profit can a criminal organization have of stealing large collections of medical data? How can they make use of them? Undoubtedly, the place where you can find information of this type put up for sale is the dark side of the Internet. Darknet, considered by some to be a mainstay of freedom and anonymity, a place where you can protect the remnants of your privacy on the Internet, by others is used for quite different purposes.

Trade of medical data in Darknet

A dark network, or Darknet, is a place that seems to the average Internet user to be an unknown space reserved for specialists in the field of com-

puter science, able to navigate through it thanks to their skills. Universal knowledge is limited to the belief that the content of this space is not only not available in search engines, but it is something more: it is based on the concept of privacy and total anonymity, the achievement of which is possible thanks to advanced cryptographic techniques. In Darknet, both Internet users and website publishers have, in principle, full protection of privacy (Omand 2015). Although large government agencies are able to track some people who function in theoretically anonymous space, it causes a number of difficulties for the services, requires a huge amount of resources and ultimately is not always effective. This triggers discussions, which are often very polarized. Defenders of complete privacy argue that granting access for law enforcement will weaken the system and create a back door for those who want to use the tool for surveillance (Jardine 2015) or other practices that violate the privacy and freedom of individuals.

It is worth to discuss here briefly how the user can gain the above mentioned anonymity. Usually, anonymity in the “dark network” is obtained using the so-called “onion network”. Typically, when you send a request, the computer automatically accesses the server on which the website you are visiting is located. When using the onion network, this direct relationship is broken, and the data runs to different servers, which in thousands of combinations “reflect them” and pass on. Thanks to this, to finally achieve the destination, our inquiry passes through many intermediaries. An effect is obtained in the form of a query that is recognized, but its sender remains unknown to the recipient, which favors anonymous communication available in most popular operating systems (Syverson, Boyce 2015). Despite the constant development of various tools enabling access to the Darknet to a wider group of recipients, the military, government institutions and law enforcement agencies are still the main users of the hidden Internet. This is because simple browsing of the Internet can reveal the user's location, and even if the content of the messages is well-encrypted, the unwanted persons can still easily come to the personal information of the communicating people and potentially discover exactly where they are. For soldiers and field agents, politicians who conduct secret negotiations and in many other circumstances that can easily be invoked, this is an unacceptable threat to security (Aschmann, Michael, Jansen van Vuuren 2017).

As results from the above, full anonymity can be achieved in the network, although often we hear that everything we do on the web can easily identify us. This is undoubtedly right, and for the average user of the network, the use of Darknet's tools is associated with something complicated, unattainable, is associated with special procedures and super-fast, modern computers. It turns out, however, that the use of the dark network is not particularly complicated – at least to some extent – and does not entail the need to invest in technological innovations. Access to hidden Internet, contrary to popular opinion, is surprisingly simple (Murray 2014). The conviction about the complicated access to this part of resources and associating Darknet with highly qualified hackers comes rather from lack of knowledge, ignorance of tools and heard, random information provided by the mainstream media. The most popular, mentioned earlier way to start the exploration of the hidden network is to install the TOR tool which means The Onion Router. It is one of the earliest and most popular onion network tools (*Darknet: ciemna strona Internetu* 2016). Technically advanced users can find a lot of different ways to configure and use TOR, but using it for the average network user without deeper programming knowledge can be as simple as installing a new browser. All you have to do is be a moderate Internet user who can install any program on the computer. Just a few clicks from the TOR site are enough to access the resources of the dark network. This browser is built on the open source code of a widely available and still popular Firefox browser, so anyone who has ever used this tool will find a browser that is friendly and easy to use (*Darknet: ciemna strona Internetu* 2016).

TOR browser gives the user a much greater potential for anonymity in the virtual space, it can be used to surf the web, while giving the user additional protection from unwanted observers of network traffic, (Chaabane, Manils, Kaafar 2010) and above all it protects the user against hackers, various forms of Internet espionage, collecting personal data without the user's will and consent, and collecting other sensitive information. The tool also gives you the opportunity to visit anonymously published websites that are inaccessible to people surfing the virtual space in a traditional way. This is one of the most frequently used browser functionalities (Spitters, Verbruggen, van Staalduin 2014).

The addresses assigned to TOR sites do not look like regular URL addresses to which the average user is accustomed. They consist of random-

looking lines of text consisting of a random string of characters. Here is an example of a hidden web address: <http://dppmfxaacucguzpc.onion/>. Connection with this type of address will move the user to the direct directory of hidden websites, but if the user does not have the required software installed, even though he or she encounters a link, after sending the request to connect to the website, no content will be displayed. In the simplest terms, the use of TOR allows you to find directories, wiki files and free links that lead the user to the content desired by the user.

It is also worth noting that TOR is not the only one of the group of tools that allows for unnoticeable passing through the resources of the hidden network. To this end, various tools and programs have been constructed so far, which are often characterized by slightly more extensive functionalities. This includes applications for sharing peer-to-peer files (direct communication between users). In this case, the most popular service, which is TOR, does not give users the comfort of exchanging p2p files, because this network has not been adapted to handle such heavy traffic (Aked 2011). The network, which is worth recalling first, is I2P (Invisible Internet Project), and the popularity of which is growing. It is a project that offers many improvements, such as integrated secure email, storage and sharing of plug-ins and integrated social features such as blogging and chat (Aked 2011). An equally popular and well-proven tool is the Freenet browser, which, like the previous one discussed above, is an anonymous peer-to-peer network available through a free download application. In this type of network, there are no centralized servers that store information or send data. Every computer that connects to the network takes the task of sharing information as it were (Lu 2016).

Therefore, as much as it is relatively difficult to break into a protected resource and illegally obtain the information there, the very procedure of offering the sale of stolen data is no longer so problematic. In principle, any user, without specialist IT knowledge, can access anonymous websites and make anonymous purchases on them. The same applies to criminals who offer large sets of sensitive data for sale. It is necessary to return to the question why the sale of medical data in the dark network is growing in popularity? What profit can be achieved by offering personalized health information?

The theft of big medical data sets is a relatively new practice. According to the data presented by T. Floyd, M. Grieco and E. F. Reid in the

article *Mining Hospital Data Breach Records: Cyber Threats to us Hospitals* (2016) in the year 2012 the first large-scale theft was recorded by the US Department of Health and Human Services. In subsequent years, there was more and more thefts on a similar scale, as a result reaching a number higher than a thousand by 2016. The analysis shows that data of various types are being stolen, in principle every medical information has a certain value. In particular, the attention of hackers focuses on the theft of personal data, dates of birth, addresses of patients, and in particular social security numbers (Floyd, et al. 2016).

Medical Health Records (EHR) offer significant benefits compared to previous systems based on traditional patient files and have been adopted by around 96% of hospitals and healthcare units in the United States (Raghupathi, Raghupathi 2014). While digitization overcomes many of the inefficiencies that have plagued medical service providers and patients themselves, the data is now more prone to burglary and theft.

In 2014, attempts were made to create a system based on artificial intelligence to combat intrusions and unauthorized access to electronic medical records. According to these assumptions, the company Proteus dealing with cyber security was established, which was to create a system available to the entire medical industry (Yao 2017). The use of artificial intelligence in a bureaucratic and risky medical environment is a challenge, but both investors and hospitals look with hope at the work on creating the system (Yao 2017). EHR contain a wealth of useful information that attracts hackers. The EHR contains all demographic data – names, historical information about the place of residence, place of work, surnames and age. There you can find family data, relatives and confidential financial information, such as credit cards and bank account numbers. We will also find there data on the history of the disease, doctor's visits, research results and even diagnosis. Medical documentation is the most comprehensive documentation on the subject of the identity of a person who is currently found in the resources of the Internet.

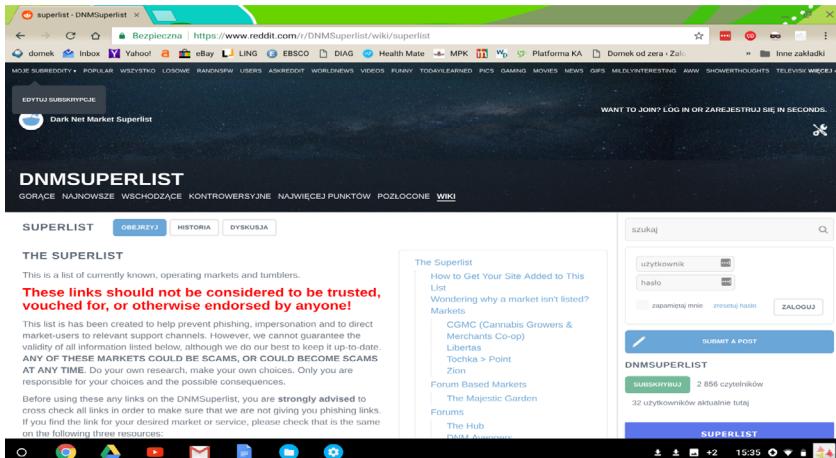
First of all, it should be emphasized that the previously described anonymity, which is ensured while surfing the dark network, creates and maintains the conditions for illegal trade in the Internet space. In 2011, the so-called Silk Road was created, which is the most popular trading venue within the TOR network. Like the popular e-Bay, Silk Road connects sellers with buyers, also offering communication

tools that allow price negotiation. The design of the sales tool enables fast and anonymous purchase of, for example, prohibited substances, various types of drugs, prescription drugs, as well as sensitive medical information. In 2013, the Silk Road was closed as a result of FBI activities, but this success did not last long, because this place was reactivated under the name Silk Road 2.0, offering the same services with even better cryptographic security (Masoni, et al. 2018).

Medical information can be worth ten times more than credit card numbers that are offered for sale in a deep network. Those who are willing to buy this type of information are mainly fraudsters who can easily use this data to create fake IDs, to buy medical equipment with reimbursement, or even drugs. You can also link the patient number to a fake service provider number or, moreover, with fictitious claims from the insurer. Unknowing patients often find that their credentials have been stolen a long time after the fraudsters used their personal medical ID to impersonate them and get health benefits (Sulleyman 2017). For example, it should be pointed out that the theft of credit card data can be quickly reported to banks that can act immediately because we have the opportunity to track the account history and relatively quickly identify unwanted transactions. Another problem is the fact that, for the time being, there is no general awareness of the meaning of personal data. The average network users do not realize how valuable the product is, and the more so that the medical documentation is one of the most prized hacker achievements. The popularity of this type of practice is enhanced by the growing strength of cryptocurrencies and strong anonymity guaranteed by Darknet.

What's even more interesting, getting to places that offer sensitive medical data in large quantities for sale does not seem particularly complicated. For a person who has knowledge of English and moderate networking skills, this is not a problem. One of the ways is to enter the Reddit forum site, a place where you can find a great majority of useful links, also to places that do not exist in the so-called surface network.

Figure 1. Reddit Superlist



Source: <https://www.reddit.com/r/DNMSuperlist/wiki/superlist>, accessed 20.01.2018.

By entering the above site, which has a simple interface, you can easily find places that offer, among others, the sale of sensitive data, including those of a medical nature. On the Reddit pages there are so-called superlists, (Reddit Superlist 2018) where you can easily find a superlist containing links to websites that are outlets for illegal data. A given superlist is a list of popular links that contain a direct reference to pages on the onion network (.onion) (Darknet). After installing a browser that allows you to connect to the TOR network in a simple way, using the links from the Superlist, you can make “purchases”, or at least, for the purpose of this study, find out what are the prices of offered, stolen data. Interestingly, although most of the data came from digital medical records, the health information alone was not popular. The vast majority of offers concerned data on the date of birth, name, place of residence, and social security number. The cost of buying such information amounted to about 1-1.5 dollars on average. However, if you assume that one theft of data means several million information about users, and that after some time some part will be sold at the above prices, the profit from such a one-off transaction can amount to several million dollars (Floyd et al. 2016)The fact that stolen medical data are more valuable than those obtained in a different way deserves special attention. This is because such theft is more difficult to detect, and the time interval for making various transactions and incurring

liabilities based on these data definitely prolongs, giving criminals time to get long-term profits. Hacking the hospital servers is therefore a lucrative occupation, because it is characterized by low costs (poor security), low risk and high, long-term rate of return (Floyd et al 2016).

Conclusions

Undoubtedly, the anonymity of Darknet's technology has two opposing sides. While people can use the hidden network for criminal purposes, they can also use it for the public good and to protect their privacy or to ensure their safety. It is also impossible to ignore the benefits of collecting and analyzing Big Data, they bring benefits in many fields, as has been shown, better and better ways of their analyzing and organizing can be a strong foundation for creating solutions that benefit the general public. However, what can work in favor is also often used in the opposite way. Therefore, there is a need for additional, more detailed investigations and analyses of the cases of theft of medical data, especially regarding the methods of obtaining them and protection against the threat of theft. The study of techniques used to manipulate the integrity of data and the ways in which data are used for healthcare frauds should be deepened, in particular by obtaining unauthorized access to medical devices. Until the collection of the above data and the construction of sealing tools for the system, hospitals and other units gathering large sets of medical data will have to continue to deal with the problem of theft, sale and various types of embezzlement based on the information obtained. According to the analysis, it can be concluded that these attacks will increase in the future, which is related to poor system security, an increasing amount of data collected and processed and, as has been shown, extremely accessible and easy to use, anonymous tools that allow data to be transferred and deriving significant financial benefits from this activity.

References:

- Aked S. (2011), *An Investigation into Darknets and the Content Available via Anonymous Peer-to-Peer File Sharing*, "Australian Information Security Management Conference", Perth Western Australia.

- Aschmann M.J., Leenen L., Jansen van Vuuren J.C. (2017), *The Utilisation of the Deep Web for Military Counter Terrorist Operations*, “Academic Conferences and publishing limited”, https://researchspace.csir.co.za/dspace/bitstream/handle/10204/9261/Aschmann_18774_2017.pdf?sequence=1, accessed 18.12.2017.
- Boyd D., Crawford K. (2012), *Critical Questions for Big Data. Provocations for a Cultural, Technological, and Scholarly Phenomenon*, “Journal of Information, Communication and Society”, 15(5), pp. 662-679.
- Chaabane A., Manils P., Kaafar M.A. (2010), *Digging Into Anonymous Traffic: A Deep Analysis of the Tor Anonymizing Network*, “Network and System Security (NSS)”, 2010 4th International Conference on. IEEE”.
- Darknet: ciemna strona Internetu* (2016), <http://www.komputerswiat.pl/artykuly/redakcyjne/2016/10/darknet-ciemna-strona-Internetu.aspx>, accessed 25.12.2017.
- Floyd T., Grieco M., Reid E.F. (2016), *Mining Hospital Data Breach Records: Cyber Threats to Us Hospitals*, in: *Intelligence And Security Informatics (ISI)*, IEEE Conference 2016.
- Gitelman L. (2011), “Raw Data” Is an Oxymoron, Cambridge-London: The MIT Press.
- Gold M.K. (ed.), *Trending: The Promises and the Challenges of Big Social Data*, in: *Debates in the Digital Humanities*, The University of Minnesota Press, Minneapolis, MN, Available at: http://www.manovich.net/DOCS/Manovich_trending_paper.pdf, accessed 28.11.2017.
- Hutchings A., Holt T.J. (2017), *The Online Stolen Data Market: Disruption and Intervention Approaches*, “Global Crime”, 18(1), pp. 11-30.
- Jardine E. (2015), *The Dark Web Dilemma: Tor, Anonymity and Online Policing*, “Global Commission on Internet Governance”, London: Chatham House.
- Krauz-Mozer B., *Czym jest teoria w politologii? Teoretyzowanie w politologii u progu XXI wieku*, https://wnpid.amu.edu.pl/~wnpid/attachments/787_Prof.%20Krauz-Mozer%20-%20referat.pdf, accessed 31.11.2017.
- Lee C.H., Yoon H.J. (2017), *Medical Big Data: Promise and Challenges*, “Kidney Research and Clinical Practice”, 36(1), pp. 3-11.
- Lessig L. (2006), *Code and Other Laws of Cyberspace*, New York, <http://codev2.cc/download+remix/Lessig-Codev2.pdf>, accessed 18.12.2017.
- Lu T. (2016), *On Privacy and Anonymity in Freenet System*, “International Journal of Security and Its Applications”, 10(5), pp. 41-56.
- Martin J. (2014), *Drugs on the Darknet: How Cryptomarkets are Transforming the Global Trade in Illicit Drugs*. London: Palgrave Macmillan.
- Masoni M., Guelfi M.R., Gensini G.F. (2018), *Darknet and Bitcoin, The Obscure and Anonymous Side of the Internet in Healthcare*, “Technology and HealthCare” 24(6), pp. 969-972.
- Murray A. (2014), *The Dark Web is Not Just For Paedophiles, Drug Dealers and Terrorists*, “Independent”, <http://www.independent.co.uk/voices/comment/the-dark-web-is-not-just-for-paedophiles-drug-dealers-and-terrorists-9920667.html>, accessed 11.01.2018.
- Omand D. (2015), *The Dark Net Policing the Internet’s Underworld*, “World Policy Journal”, 32(4), pp. 75-82.

- Pang S. (2017), *Malicious Events Grouping via Behavior Based Darknet Traffic Flow Analysis*, "Wireless Personal Communications", 96(4), pp. 5335-5353.
- Raghupathi W., Raghupathi V. (2014), *Big Data Analytics in Healthcare: Promise and Potential*, "Health information science and systems", 2(1)/3, doi:10.1186/2047-2501-2-3.
- Rosa A., Chen L.Y., Binder W. (2014), *Predicting and Mitigating Jobs Failures in Big Data Clusters*, in: *Cluster, Cloud and Grid Computing (CCGrid)*, "2015 15th IEEE/ACM International Symposium".
- Sutton M. (2016), *Monitoring Darknet Access to Identify Malicious Activity*, "U.S. Patent", <https://patents.google.com/patent/US8413238B1/en>, accessed 10.12.2017.
- Sulleyman A. (2017), *NHS Cyber Attack. Why Stolen Medical Information is so Much More Valuable Than Financial Data*, "The Independent", <http://www.independent.co.uk/life-style/gadgets-and-tech/news/nhs-cyber-attack-medical-data-records-stolen-why-so-valuable-to-sell-financial-a7733171.htm>, accessed 11.01.2018.
- Syverson P., Boyce G. (2015), *Genuine Onion: Simple, Fast, Flexible, and Cheap Website Authentication*, <https://www.nrl.navy.mil/itd/chacs/sites/www.nrl.navy.mil.itd.chacs/files/pdfs/15-1231-0478.pdf>, accessed 04.01.2018.
- Van Buskirk J. (2016), *Characterising Dark Net Marketplace Purchasers in a Sample of Regular Psychostimulant Users*, "International Journal of Drug Policy" nr 35, DOI: <https://doi.org/10.1016/j.drugpo.2016.01.010>.
- Wehinger F. (2011), *The Dark Net: Self-Regulation Dynamics of Illegal Online Markets for Identities and Related Services*, "Intelligence and Security Informatics Conference (EISIC)". IEEE.
- Wolniak A. (2006), *Kod, normy społeczne, prawo i rynek: cztery regulatory zachowań w internecie i ich wzajemne zależności*, Centrum Badań Problemów Prawnych i Ekonomicznych Komunikacji Elektronicznej, Uniwersytet Wrocławski, <https://app.lumnpdf.com/viewer/8c5Q5aAqBKNj7mikN>, accessed 04.01.2018.
- Yao M. (2017), *Your Electronic Medical Records Could Be Worth \$1000 To Hackers*, "Forbes", <https://www.forbes.com/sites/mariyayao/2017/04/14/your-electronic-medical-records-can-be-worth-1000-to-hackers>, accessed 04.01.2018.

(Big) Data Activism – Trends, Challenges and Opportunities

Abstract: The article's purpose is to introduce the subject of data activism which, related to the development of Big Data technology, offers activists new possibilities, but also confronts them with a number of new challenges. The author briefly presents Stefania Milan's typology of data activism, which distinguishes two types: *reactive* and *proactive collective action*, and then proposes her own version of this typology, comprised of three main currents of action: *anti-data activism*, *pro-data activism* and *inter-data activism*, providing examples of both technological and educational actions in each of these areas.

Keywords: data activism, Big Data, open data, privacy, anonymity, surveillance, transparency

Introduction

In 2012, Alexis C. Madrigal analyzed the changes that arise from within activism, especially in the field of culture jamming, consisting, among other activities, of the subversive disruption of corporate advertising. He pointed to the challenges and opportunities of operating in a digital environment. He predicted that the best ways to fighting against the corporate message would be deliberate (subversive) interaction with displayed advertisements, and the disruption of algorithms of data collected about Internet users. For example, in an arranged attack, information is sought on a particular, determined subject by many Internet users, clouding the data picture that is generated on this basis. Madrigal compared such a subversive search for information to the well-known “hactivist”

distributed-denial-of-service (DDoS) attacks organized by Anonymous. If collective action were on a suitably large scale, it could influence the quality of the data collected by, for example, Google. He concludes by asking when activists will understand that if they act in such a way that data is more difficult to analyse, they make use of an important political tool (Madrigal 2012). Not only the obstruction of the analysis and collection of data, and thus strictly resistance action, but also positive action – the acquisition and use of small as well as Big Data, especially from algorithms based on Big Data – constitutes an important element of data activism.

The aim of this article is to introduce the subject of data activism, defined as a form of digital intervention, or broadly digital mobilisation for social and political change with the use of (big) data. It embraces various activities devoted to both data protection and disclosure, to the use of data in organizing actions and activist campaigns, but also to education for data. The data in this regard is considered a direct objective of the intervention and means to act.

Data activism, related to the development of Big Data technology, offers activists new opportunities, but also confronts them with a number of new challenges. Its reflections on data activism have been structured in a way that is consistent with the typology that I have elaborated in the text. I distinguish three types of data activism: *pro-data activism*, *anti-data activism* and *inter-data activism*. Examples of activist practices will be used to illustrate each of these described types. Finally, I will indicate the need for an extensive educational campaign, exemplified by the activities of the School of Data.

Data activism typology

The question that Madrigal asked in 2012 still seems valid, although undoubtedly activists are increasingly aware of the opportunities and threats arising from the use of data by governments, corporations and the activists themselves. However, much remains to be done in response to this question, both its theoretical and practical aspects. Maya I. Ganesh and Morana Miljanovic, of the Tactical Technology Collective,¹ pointed out

¹ The Tactical Technology Collective is an international non-profit organisation which for years has been raising awareness of issues related to privacy, functioning in the digital world, and using information to take effective action. It does this by working out practical solutions and through research, training and workshops. (Tactical Technology Collective, no date, a).

that it was essential for activists to “co-operate” with data. In their short text, “A small list about Big Data”, they attempted to acquaint activists with the subject of “Big Data” from various perspectives, encouraging deeper reflection about it, and especially an answer to the question: Why should activists worry about Big Data? The authors give three main reasons:

1. surveillance,
2. prediction,
3. transparency.

Ad 1. Ganesh i Miljanovic emphasise that Big Data technology perpetuates mass surveillance, as – among other reasons – almost the whole technological infrastructure, software and the law that governs them, are subject to orders from the U.S. government, through telecommunications firms, in co-operation with their British, Canadian, Australian and New Zealand partners.

Ad 2. As a positive example of the use of Big Data technology activism, they note the possibility of predicting outbreaks of epidemics, unemployment and the results of natural disasters.

Ad 3. The third reason, essential for activists, is transparency, which Big Data technology makes possible. The authors note, all the same, that this may be a doubtful advantage, in practice resulting in breaches of privacy (Ganesh, Miljanovic 2014).

As the authors also emphasise, these three reasons are not irrelevant to one another, as there are obvious tensions arising from initiatives to increase data transparency, when they brush against growing concern for the right to privacy and anonymity.

Data activism according to Stefania Milan

More complex and scientific evidence, and answers to the question about the significance of (big) data for activism, may be provided by Stefania Milan’s project, “Dataactive. The Politics of Data According to Civil Society”. Financed by the Starting Grant of the European Research Council, it focuses on three main research questions: *How do citizens resist massive data collection by means of technical fixes? How do social movements use Big Data to foster social change? How does data activism affect the dynam-*

ics of transnational civil society, and transnational advocacy networks in particular? (The Research Questions, DATACTIVE, n. d.)

The project that Milan leads assumes a multidisciplinary approach, making use both of qualitative and quantitative methods. It is intended to make it possible to analyse forms of organization, repertoires of activity and the role of programming in data activism, and to identify emerging structures and transnational strategies of support networks (transnational advocacy networks).

Milan, proposing an interdisciplinary perspective on research into data activism, refers to the following fields of research:

- a) political sociology, and especially social movement studies, to understand the dynamic nature of collective action, organizational strategy and sense-making activities;
- b) media studies, to show the communication processes within data activism, as well as their correlation with journalism, and with alternative and social media;
- c) international relations, to follow changes in authority within transnational civil society;
- d) software and platform studies, for their technological support;
- e) science and technology studies, as they promote, simultaneously, thinking in terms of “sociologism” and “technologism” (Milan 2017, 158).

In Milan's view, data activists draw from the traditions of hacking and open-source movements, as well as the radical tech activism of the 1990s, statistical activism (statactivism) and the “do-it-yourself” cultures of hacklabs and hackerspaces (Milan 2017, 155). To this list, the subversive tradition of culture jamming may be added.

At this point, the definition of data activism which Milan adopted bears quoting. According to this definition, it is the most recent form of media activism, “a rapidly growing phenomenon at the intersection of the social and technological dimension of human action”: technological, as it functions in a technical environment, and socio-political, because it is developing in the context of socio-political processes and seeks to change the distribution of power.

Milan distinguishes two types of data activism: *reactive* and *proactive*

collective action. Reactive activism refers to situations where citizens put up resistance to works that threaten their civil rights, especially their right to privacy, which may be the effect of corporate or government surveillance. They fight, then, in various ways against the “surrender” of their own small data, and against the possibility that it can be processed in Big Data.

Proactive activism makes use of the opportunities that Big Data offers to organize supportive actions and campaigns – that is, uses them to promote social change (Milan 2018, 153-156; Milan, Guetierrez 2015, 120-123).

A suggested functional typology

Milan's typology is based on opposition. I propose to clarify this typology, referring as well to Ganesh and Miljanovic's postulates, taking into account the function of data in activists' work. I will thus distinguish three types of activity involving data, arising from the purpose and methods of this work:

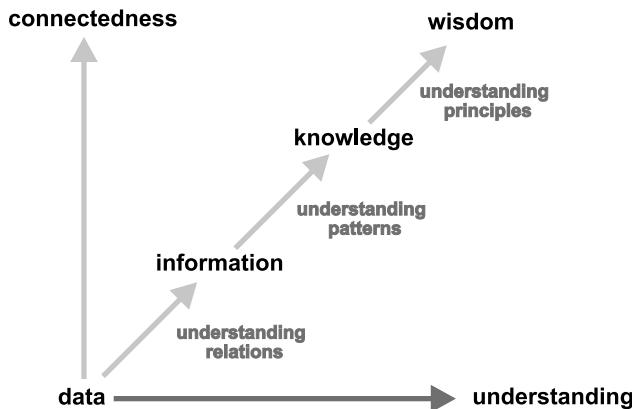
1. activity to release data (**acting for**)
2. activity against data (**acting against**)
3. activity using data (**acting by**).

I distinguish three types of data activism, corresponding to these types of activism:

1. **anti-data activism**, which would define activism intended to protect people from data becoming available; thus this activity aims at the anonymity of messages and participants in communication, and generally at protecting data from corporate and government surveillance. Thus it would be, according to Milan, *reactive collective action*.
2. **pro-data activism**, which would define actions which, in particular, make data public, freeing it from various closed interests (financial or political).
3. **inter-data activism**, which would define actions and campaigns that use data to organize, set goals, build arguments, and generally promote social change. This type of activism is based on interaction with data in devising a message and – speaking broadly – strategies of actions; Milan would locate *inter-data activism* in the area of proactive data activism.

In this context, it is also worth examining a version of the pyramid of knowledge proposed by Bellinger, Castro and Mills (no date).

Figure 1. Transition from data to wisdom



Source: Bellinger G., Castro D., Mills A. (no date), *Data, Information, Knowledge, and Wisdom*, <http://www.systems-thinking.org/dikw.htm>, accessed 01.02.2018.²

In this schema, data is at the lowest level; information stands higher. If we assume that data, as opposed to information, is raw and unanalysed, while information is data that has been analysed and placed in a context, then it may be concluded that in pure data activism *anti-data activism* and *pro-data activism* may be recognised, where we are fighting either against making data public, or in favour of data. *Inter-data activism* moves data to another level of the pyramid, namely the level of information. When data is embedded in a context, where it is analysed and relationships among items of data are pointed out, it is no longer mere data, but has been transformed into information. Because of such a view of data, we are not speaking any longer of *data activism*, but of *info-activism*. As a result, it may be worthwhile treating (big) data activism not only in the way that Milan proposed from a technological perspective, as the latest form of media activism, but also as the basis of info-activism³.

² Bereziński, Hołubiec i Wagner points out that this schema is not entirely correct, as it may be assumed that there are links in this chain where thought is suspended (Bereziński, Hołubiec, Wagner 2009, 11).

³ I write more on the concept of *info-activism* in: *Info-Activism: The egalitarisation of Access to Information*, (2016).

Pro-Data Activism

Pro-data activism, or activism that aims to fight for information to be freed or made public, is to a large extent associated with the open data movement (linked to initiatives such as Open Knowledge and Open Governance). This movement, as Baack emphasises (2015, 2), blends two mutually opposed transformations of knowledge and power: datafication and the proliferation of the hacker culture, especially the open source culture. On the one hand, activists focus on the distribution and use of data related to the datafication and quantification of community life, whose strongest expression is Big Data. On the other hand, it is especially open data activists who draw from the open-source culture in the creation and use of data. Baack points out three main practises and ideas of open data activists:

1. Through the use of data as a necessary condition for generating knowledge, activists change the idea of using source code to share raw data, allowing others to create their own interpretations and to generate their own knowledge. This in turn becomes the expression of the democratisation of information. In this, a fundamental reservation becomes evident, related to the concept of “raw data”, which has been criticized by, for example, L. Gitelman and V. Jackson. They emphasise in the introduction of their work, significantly titled *“Raw Data” is an Oxymoron*, that: *At first glance data are apparently before the fact: they are the starting point for what we know, who we are, and how we communicate. This shared sense of starting with data often leads to an unnoticed assumption that data are transparent, that information is self-evident, the fundamental stuff of truth itself. If we’re not careful, in other words, our zeal for more and more data can become a faith in their neutrality and autonomy, their objectivity.* (Gitelman, Jackson 2013, 2-3).

In absolute terms, it is indeed difficult to talk about raw data, as the algorithm for collecting it imposes specific interpretative frames. For activists, however, data exist in a specific perspective: “raw” means, to them, “as collected”.

2. If activists recognise information as a pre-condition of political participation, they associate it with open and elastic forms of repre-

sentative democracy, implementing the participatory open source model in the context of a strengthened role for citizens in decision-making proces and in local community activity. Such a model demands a change from a less authoritarian relationship between authorities and citizens, to one which is more collaborative, open and flexible. It will be open because it results from greater transparency, through the sharing of raw data and opennesss to the role that citizens play in decision-making processes. It will be elastic because it is adapted to current needs and a local context.

3. According to activists, raw data must be refined in order to transform it into knowledge that is accessible to citizens. As a result, activists attempt to create or become “data intermediaries”, co-operating with professional journalists or becoming themselves programmer-journalists, or creating civic technologies. Baack points out three essential features of data-empowering interme-diaries, according to activists. They should be (a) data driven, able to work with large and complex data sets, so that they are available to others; (b) open, creating data from which a story may be generated, or create applications that are available to their public; and (c) engaging, actively getting citizens involved in public affairs (Baack 2015, 3-7).

The model presented in Baack’s research seems ideal and well grounded in theory, on the one hand; on the other, it demands that activists be unusually ethical and very often and paradoxically, both transparent and invisible, especially when it comes to the interpretation of data. At every stage of an activity, many possibilities of ethical conflict are possible. Much as the first stage seems the most neutral, in spite of reservations related to the influence of algorithms, it is still here where distortions related to the selection of data may arise; not only small amounts of data, but also Big Data. At further stages, where data is embedded in a context and transformed into information, then set into an interpretative framework, objectivity is demanded, not only of theory-based and often complex knowledge, but also of that which is technological, political and social. Just as necessary are the skills of presenting data so that it is accessible and adapted to its audience, in ways that are accessible and adapted to the level of the public. Objections to

open data initiatives are, on the one hand, a consequence of the opposition of the right to openness to the right to privacy; and on the other, arise from an underlying problem in digitalisation, known as “digital divide” or “digital exclusion”. In this context, it may be known as “data divide” or “data exclusion”.

Among those who have formulated critical and well-grounded questions and reservations related to the open data phenomenon is M. Gurstein (2011). He has listed seven factors essential for open data to function so that it is in fact also accessible including to marginalised groups and not acting, as some examples show, against them; in other words, where essential elements are present that can counteract the data divide and which are essential for end users to make effect use of open data. These factors are:

1. The Internet: accessibility to Internet with suitable capacity; physical accessibility for the handicapped; accessibility to networks in spite of attempts, such as political efforts, to limit it.
2. Computers and software: accessibility to suitably powerful computers and programmes that can carry out essential analysis.
3. Computer/software skills: the essential knowledge and skills for using programmes and making data available in a visual form, so that it is accessible to typical users and not only specialists.
4. Content and formatting: data presented in suitable formats in terms of language, coding, and analysis, as well as access (not economically limited) to professionals who can ensure this.
5. Interpretation/sense making: the ability to select data and knowledge and place it into a context, including a local context, thus transforming data into information.
6. Advocacy: social and individual support; people and groups capable of transforming data so that actions on behalf of local communities are more effective.
7. Governance – an appropriate financial, legal, regulatory and political system, allowing the efficient use of data.

It is difficult to disagree with the author, as Big Data technology is now being “swallowed” in such a way that the basic conditions for en-

suring user equality may be forgotten. Unfortunately, this equality will remain only theoretical: digital base conditions should be supported just as other primary needs are satisfied, which is difficult for the poorest and marginalized groups. So activists play an even bigger part in creating suitable conditions for educating others in Big Data technology – in a technological context, as well as social and political.

Anti-Data Activism

Anti-data activism, fighting for the right to privacy, is nothing new in activist practice, whether in the analogue or digital worlds. Here, the well known and documented actions of the Hactivismo group are worth mentioning, and their projects to make messages anonymous (the web browser TorPark; ScatterChat, a safe way for dissidents to communicate, as whole messages are encrypted; and Camera Shy, an application which makes it possible to hide data in graphic; and other forms).

Very practical tools include, for example, Freedombox, free software that powers a personal server, acting as a protective shield between the user and the public Internet. It is easy to use and functions as well as an access point for Wi-Fi. Freedombox is equipped with a private browser, allowing messages and voice connections to be encrypted. It stores user data securely, protecting its anonymity (Freedombox Foundation, no date; see also Taffel 2014, 65).

Big Data technologies have taken this type of activism and surveillance to a completely different level, but it is still often unavailable to users of digital technologies⁴ due to lack of awareness. That is why, in this type of activism, not only technology is important, but above all education in data activism, focused on the protection of the right to privacy.

One of the most interesting initiatives is certainly *The Glass Room* exhibition, organised by the Tactical Technology Collective. This group, in co-operation with Mozilla, organized *The Glass Room London: Looking into your online life*⁵ exhibition at the end of October and in early

⁴ This is in spite of the efforts of, for example, E. Snowden, who made public information about the PRISM system of massive citizen surveillance.

⁵ This was the second edition of *The Glass Room*. The first, *The Glass Room New York City*, was organized there in 2016.

November, 2017. It was combined with workshops, panel discussions and lectures, where the main subject was data. The exhibits were divided into four groups, with telling titles: *Something to hide*, *We know you*, *Big Mother*, *Open the box*. The first group (*Something to hide*) presented, often provocatively, how various companies that collect our data may use it. The idea of this part of the exhibition was to make us aware that even when we are guided by the conviction that we have nothing to hide, it is helpful to understand what we do not hide. Mention should be made, here, of Sam Lavigne's *Online Shopping Center*. He designed a brain wave scanner that recorded brain activity while subjects were thinking of shopping, and while they were thinking of their own mortality. The result of this work was an algorithm that recognised two mental states: "shopping-like" and "death-like". Lavigne put the scanner on during sleep, connecting it to the computer, which then did shopping. The viewers could put on the scanner and check if the algorithm recognised their mental state.

Another activist artist, Manuel Beltrán, researched the way in which the data we create constitute a form of productive work. His *Data Production Labor*, recording the movements of subjects' hands over a period of time while they were scrolling Facebook, and software serving to recognise emotion, found correlations between the reading of content and emotions, and visualized the user's contribution to the Big Data economy. Further works showed, for example, how data can sound or smell.

The group of exhibits *We know you* concentrated on privacy, on the question of how technological companies are rooted in the lives of typical citizens, and how much they know about them. A most perverse exhibit was a model of Mark Zuckerberg's house, produced on the basis of pictures taken from Google Maps, covered by a glass dome, which represents an incident where Zuckerberg, to ensure his right to privacy, purchased four houses that neighboured his own. Privacy and data protection have become a remarkably expensive life investment that many will not be able to afford.

The *Big Mother* section presented public surveillance camouflaged as "caregiving" applications, such as "Silver Mother™", providing care for older relatives; "Iris Scanning", used by the United National High Commission for Refugees to scanning the irises in the eyes of refugees, allowing them to withdraw money from banks and automatic transfer

machines, but also making it possible to track them; and films about the firm 23andMe, involved in DNA research and providing reports on the history of a family and the risk of illness, but at the same time earning money by selling access to the data sets generated on the basis of these tests. The *Open the Box* section served to raise viewers' awareness of the use of data during elections ("Data and Elections"), on dating sites ("Data and Dating") and at home ("Data and Home"): devices creating the "Internet of things" and data collected about residents.

At the centre of the exhibition was the Data Detox Bar, where visitors could speak about projects, obtain individual advice about Internet privacy and regain some control over their data in the digital world (The Glass Room, no date). As the remarks and reflections left by the visitors show, the exhibition was remarkably necessary, as it made the visitors conscious of operations in the digital world about which they usually were unaware. As the organizers not only raised this consciousness, but also provided education about this data, many of the guests stated that they would change their digital lives (Tactical Technology Collective, no date, b)

Inter-Data Activism

An example of this type of activism is the well-known crowdsourcing of data, for example, through the help of the open-source platform Ushahidi, which made possible to collect grassroots data, and then map it. Ushahidi (in Swahili, "witness"), was used as early as in 2008 in Kenya to provide a visual representation of violent attacks during post-election rioting. Ushahidi collects information by e-mail, the Internet, sms-es and mms-es. It makes possible the creation of activist maps of crises, whether political, social or of natural catastrophes. Ushahidi, which P. Meier considers a liberation technology, has been used many times as a tool of the citizen society to monitor elections in Afghanistan, Burundi, Egypt, India, Kenya, Kyrgyzstan, Lebanon, the Philippines, Mexico, Mozambique, Sudan and Tanzania (Meier 2012, 95-98). It was also used during the times of earthquakes in Haiti and Chile, fires in Russia, and floods in Australia and the United States. Interestingly, it was also effective during the floods in Poland in 2010. The Pretius firm was asked by the Stocznia Foundation

for help in filling its technological gap, involving the provision and mapping of data about threats, assistance needs for disaster victims, and so on. Programmers took two products into consideration: Sahana and Ushahidi. Neither completely met their expectations, but the Ushahidi version proved most responsive to the demands of responding to needs during the earthquake in Haiti. The programmers made essential improvements and extensions to the source code. After the map was uploaded to the address <http://alert.powodz.ngo.pl>, the number of visitors grew each day, and one of the most-often used functions was the possibility of subscribing to alerts. Many people in need have received essential help from other users following situations, thanks to the very help of the platform created by Pretius programmers (Olejnik 2010).

Information crowdsourcing is often used in mapping incidents of sexual harassment in the public space. Apart from Ushahidi, used in Cairo by the Harrasmap organisation; similar safety applications in the Middle East are, for example, Ramallah Street Watch, Safe Streets Yemen and Resist Harassment Lebanon (Bramley 2015). One may also point to the SAFECITY platform, used in India, where a rape takes place every 20 minutes and 95% of women do not feel safe in public. The SAFECITY platform collects personal histories of sexual harassment in public areas. The collected data allows trends to be identified, and maps of dangerous areas to be created (TechRasam 2017).

As the above examples show, when bottom-up data is collected and presented in a visual form, it has a substantial influence on the functioning of a local community. Women facing harassment can avoid unsafe places; during natural disasters, maps can be made of places where help is most needed; and at the same time such help, even non-institutional, can be provided, though within the local community. What is especially important in the context of the above actions is the ease of use of the application; they do not require specialised skills, which often presents problems in the effective use of data.

Data education

In the context of data activism, educational work at all levels seems particularly important. The first stage in working with data entails making

people aware of data, of opportunities to gain data from users and of the ways that they use data. A second necessity is technological education, so that not only specialists and persons who are already privileged in different respects may use data, often against marginalised groups. It is essential that action be taken to wipe out digital illiteracy, and thus data illiteracy, as well. A perfect example of such activity is found in the effort that the School of Data has made: this organisation arose in 2012 as the effect of the co-operation of Open Knowledge International and the Peer-to-Peer University.

The School of Data is guided by the following key values:

- empowerment: the strengthening of individuals and organisations by equipping them with the skills of working with data (data literacy), which would allow them to work better in the environment, understand it better and in consequence achieve real changes;
- collaborative: “learning by doing” and “work with real data”, two perspectives, fulfilled in the course of educational events (such as at summer camp holidays), and targeted workshops;
- responsible and ethical use of data: this is linked to setting limits on the use of data, and to questions of privacy and supervision;
- evidence is power: promoting a dialogue based on data and evidence;
- non-partisanship: the organisation does not support any political option (School of Data, no date, a).

The School of Data bases its activities on three main programmes: the *School of Data Fellowship*, the *School of Data Network Lab* and *School of Data Curriculum Development*.

The first of these programmes is flagship programme, offering ten-month grants to data practitioners or enthusiasts. The purpose of the grant is to broaden awareness and skills of working with data, as well as to build communities which, thanks to data literacy, may conduct and initiate required changes.

The School of Data Network Lab allows pilot projects involving data to be brought to fruition, and thanks to small grants, to disseminate knowledge about data in local communities.

The School of Data Curriculum Development: from the beginning of the School of Data's existence, the curriculum has been the main focus

of its activity. It was developed during workshops and at camps, by volunteers and students on grants, and partners throughout the world. The current (multi-modular) programmes, operating in ten languages and available on the Internet page, explore a variety of aspects of functioning in the world of data. The whole programme consists of the following packages:

- a) *Data Fundamentals*: an introduction to the subject of data, pointing out the main sources of data, on the basis of ways of representing it visually.
- b) *Data Clearing*: created by the Tactical Technology Collective, this module focuses on the importance of cleaning data, and how to do this.
- c) *Exploring Data*: prepared by the Tactical Technology Collective and the School of Data, this teaches how data may be understood, and how to use data spreadsheets and SQL language for data analysis.
- d) *Extracting Data*: on extracting data from the Internet, and maintaining it in a machine-readable formats.
- e) *Mapping*: an introduction to on-line geocoding and the Quantum Geographical Information System (QGIS), open geoinformation software enabling the collection, analysis, processing and sharing of spatial data, and the creation of maps.
- f) *Collecting Data*: devoted to the collection of data using smart-phones, this course presents, among other tools, the Open Data Kit, an open tool for creating questionnaires, gathering data and building databases, collecting them on the server and presenting them in useful formats.
- g) *Presenting Data*: devoted to various methods of presenting data in a visual form.
- h) There are other thematic courses, such as *An Introduction to Aid Data*, focusing on various aspects of using data related to international aid (School of Data, no date, b)

The educational programme is well-considered, extensive and mul-

ti-faceted. It provides for the acquisition of a number of basic technical skills on the one hand, and on the other, creates an awareness of data: its possibilities and ways of using it for the common good, without overlooking the problems related to surveillance and privacy protection.

Conclusions

Data education is an essential condition of the success of data activism. Projects such as the School of Data, HaktionLab, and the Tactical Technology Collective are important in building awareness of data, on the one hand, and on the other to equip activists, local community organisers and typical computer users with the skills related to the gathering, use and protection of data, according to the proposed typology: *pro-data activism*, *inter-data activism* and *anti-data activism*. A further stream of action about data is the creation of accessible, easily used programmes that make it possible to work with data, and which do not require very expensive equipment, so that data is accessible – a necessary condition of closing the data divide and preventing data exclusion. Data activism activities will serve to ensure that data remains accessible (and not only to privileged groups); understandable, thus able to be processed and analysed; interpretable, able to be placed in an interpretative framework; and effective, used for local community needs to bring about desirable changes.

References:

- Baack S. (2015), *Datafication and Empowerment: How the Open Data Movement Re-Articulates Notions of Democracy, Participation, and Journalism*, “Big Data & Society”, July 22.
- Bellinger G., Castro D., Mills A. (no date), *Data, Information, Knowledge, and Wisdom*, <http://www.systems-thinking.org/dikw/dikw.htm>, accessed 1.02.2018.
- Bereziński M., Hołubiec J., Wagner D. (2009), *Hierarchiczna struktura poznania – piramida wiedzy*, pp. 5–17.
- Bramley E. (2015), *Can the Safecity App Make Delhi Safe for Women*, “The Guardian”, 13.08, <https://www.theguardian.com/cities/2015/aug/13/can-the-safecity-app-make-delhi-safer-for-women>, accessed 1.02.2018.

- DATACTIVE (no date), *The Politics of Data According to Civil Society, The Research Questions* <https://data-activism.net/about/>, accessed 15.01.2018.
- Dytman-Stasieńko A. (2016), *Info-Activism: The Egalitarisation of Access to Information*, in: Konarska K., Szymol A. (eds.), *Media and Journalism in the Digital Era*, Wrocław: Wydawnictwo UW, pp. 119-130.
- FreedomBox Foundation (no date), *About the Foundation*, <https://freedomboxfoundation.org/>, accessed 1.02.2018.
- Ganesh M.I., Miljanovic M. (2014), *The Small List About Big Data*, <https://visualisingadvocacy.org/blog/small-list-about-big-data>, accessed 1.02.2018.
- Gurstein M. (2011), *Open Data: Empowering The Empowered or Effective Data Use for Everyone*, "First Monday" 2-7.02, <http://firstmonday.org/ojs/index.php/fm/article/view/3316/2764>, accessed 1.02.2018.
- HaktionLab a (no date), *Welcome to the HaktionLab*, https://hacktionlab.org/Welcome_to_HacktionLab, accessed 1.02.2018.
- HaktionLab b (no date), *Tech Tools for Activists*, https://hacktionlab.org/Welcome_to_HacktionLab, accessed 1.02.2018.
- Gitelman L., Jackson V. (2013), *Introduction*, in: Gitelman L. (ed.), (2013), "Raw Data" is an Oxymoron, Cambridge-London: The MIT Press.
- Madragalis A.C. (2012), *The New Culture Jamming: How Activists Will Respond to Online Advertising*, "The Atlantic", 12 May, <https://www.theatlantic.com/technology/archive/2012/05/the-new-culture-jamming-how-activists-will-respond-to-online-advertising/257176/>, accessed 1.02.2018.
- Meier P. (2012), *Ushahidi as a Liberation Technology*, in: Diamond L., Plattner M.F. (eds.), *Liberation Technology: Social Media and the Struggle for Democracy*, Baltimore: The John Hopkins University Press, pp. 95-109.
- Milan S. (2017), *The New Frontier of Media Activism*, in: Pickard V., Yang G. (eds.), *Media Activism in the Digital Age*, London-New York : Routledge, pp. 151-163.
- Milan S., Gutiérrez M. (2015), *Citizens' Media Meets Big Data: The Emergence of Media Activism*, "Mediaviones", nr 14, Enero-Junio, pp. 120-133.
- Milan S., van der Velden L. (2016), *The Alternative Epistemologies of Data Activism*, http://digicults.org/files/2018/01/Stefania-Milan-Lonneke-van-der-Velden_Data-activism.pdf; accessed 1.02.2018.
- Olejnik R. (2010), *Poland: How the Ushahidi Platform Was Used to Respond to Flooding* (guest blogpost), <https://www.ushahidi.com/blog/2010/12/23/poland-how-the-ushahidi-platform-was-used-to-respond-to-flooding>, accessed 1.02.2018.
- School of Data a (no date), *Our History, Our Core Values*, <https://schoolofdata.org/our-history/>, accessed 1.02.2018.

School of Data b (no date), *Programmes*, <https://schoolofdata.org/programmes/>, accessed 1.02.2018.

Taffel S. (2014), *We Have Never Been Open: Activism and Cryptography in Surveillance Societies*, “MEDIANZ”, Vol. 14 No. 2, pp. 52-72.

Tactical Technology Collective (no date) a, *About us*, <https://tacticaltech.org/pages/about-us/>, accessed 1.02.2018.

Tactical Technology Collective (no date) b, *Public Response to The Glass Room London*, <https://tacticaltech.org/media/itl/dec17/release-glass-room-public-response.pdf>, accessed 1.02.2018.

TechRasam (2017), *How Online Crowdsourced Data is Helping Safe City Understand Patterns of Sexual Violence in India*, <https://techrasm.org/2017/07/14/safecity/>, accessed 1.02.2018.

The Glass Room – Looking Into Your Online Life (no date), <https://theglassroom.org/>, accessed 1.02.2018.

Data Journalism. Essence, Technology, Development

Abstract: The following paper examines data journalism, and it aims to show its essence and the technologies through which it operates, and to track its development. Data journalism is one of the vastest opportunities provided by online media for the way media products are created, for their types, and for the investigative journalism. The paper examines how data journalism emerges as a professional practice and the first scientific research studies about it. The conclusions are that data journalism redefines online media, investigative journalism and it requires specific new journalistic competences in order to be used in professional media.

Keywords: data, journalism, big data, data visualization, journalism skills

Introduction

The development of digital technologies, means of communication and software products leads to a number of transformations in media that are functioning and evolving in an online environment. The hypertext nature of the Internet provides a number of new opportunities and forms of journalism as one of them is data journalism.

Combining the massive increase in the amount of data on the Internet with the technologies, through which they can be detected, processed, analyzed and transformed into media products, lies at the heart of this phenomenon in online media, typical for the second decade of the 21st century.

The center of this paper is data journalism, and the subject is to determine its essence, the technologies through which it operates, and to track its development.

The goals are to study a media phenomenon which is still paving its way not only in Bulgarian scientific literature, but also in the practical media environment; to outline definitions; to examine the essence of data journalism; and to highlight the most important moments of its development. In order to achieve that, research tasks will be carried out to analyze Bulgarian and foreign literature on the topic, to analyze case studies and manuals as well as results of previous research of the author, related to the topic.

Data journalism is one of the vastest opportunities provided by online media for the way media products are created, for their types, and for the investigative journalism. Data journalism is an entirely new media phenomenon that is yet to undergo its rapid development, whereas trends in global technology and media are a promise of its evergrowing future influence.

What is data journalism?

From an academic point of view, the beginning of data journalism is set out by Jonathan Hewett (2016), who taught it for the first time as a subject at City University London. Another pioneer of data journalism is Simon Rogers of “The Guardian”, who forms a dedicated team for this type of journalism in the paper where he works and has laid its practical foundation (*The Guardian Data Journalism Blog*). He argues that data journalism can be practiced by anyone with a computer and the necessary skills, but that it still relies on serious funding from major media brands (O'Reilly 2012).

Journalism has always used data and information to fulfill its duty. But then what distinguishes data journalism from traditional journalism? The Web holds a vast amount of data that cannot be physically reviewed and analyzed. The volume is now measured with the unit Zettabyte (ZB). The amount of digital information contained in a Zettabyte is impressive – the number 10 raised to the 21st power (10^{21}) or about 1 000 000 000 000 000 000 bytes. The equivalent in more widespread units would be one billion terabytes or one trillion gigabytes.

The essence of data journalism is to find data, process it, capture important information, make the necessary conclusions, and visualize the result using computer technology in such a way as to “translate” to the audience the complex content of the databases it uses, and to help said audience understand it.

Data is getting more and more accessible on the Internet – from publicly advertised data, that public administrations are required to publish, to “leaked” secret data that instantly focuses our attention on examples like WikiLeaks where more than 300 000 documents have been published relating to the US Army’s actions in the wars in Iraq and Afghanistan. In the summer of 2016, nearly just as many emails were published, containing pieces of communication between members of the ruling party in Turkey. Data on US military actions has helped journalists visualize the content of all documents into projects that have now become textbook examples, related to the practice of data journalism (Stray 2010).

One of the challenges before data journalism is that often information available to professional journalists is in an inappropriate file format, such as scanned PDF documents or various databases, presented in different units of measurement, currencies, complex specific abbreviations and encoded messages. For the end product to be reliable and of high quality, everything necessary needs to be extracted, systematized, unified, and only then analyzed in order to yield the desired information. Journalists need to know what they are looking for and how to “ask” the data, how to “interview” the available data in order to get the desired answer (Rogers 2012). After that comes the preparation of visualization, which should be as clear and user-friendly as possible so that it can be understood by the average audience representative.

Data journalism provides journalists with entirely new horizons. With free software tools and accessible information, they can carry out their professional mission in a high quality way – working for the benefit of society. Data journalism gives an explanation of events, reveals trends in societies and countries, turning “raw” data into something enriching and easy to explain. The World Wide Web creator, Tim Berners-Lee, has an interesting definition of data journalism: *Data journalism is the future. Journalists need to be data-savvy. It used to be that you would get stories by chatting to people in bars, and it still might be that you'll do that*

way sometimes. But now it's also going to be about poring over data and equipping yourself with the tools to analyse it and pick out what's interesting. And keeping it in perspective, helping people out by really seeing where it all fits together, and what's going on in the country (Gray, Bounegru, Chambers 2012, 6). Data journalism can be used not only for political and economic purposes, but it could also find its place in other areas of journalism – entertainment, sports, culture, etc.

Other definitions by university lecturers and professional journalists to help understand the essence and importance of this type of journalism are as follows:

- “Filtering the Flow of Data”, prof. Philippe Meyer, *University of North Carolina*;
- “New Approaches to Storytelling”, Aaron Phillhopper, *New York Times*;
- “Like Photo Journalism with a Laptop”, Brian Boyer, *Chicago Tribune*;
- “Filling the Gap Between Statistics and Narrative,” David Anderson, a freelance journalist;
- “A Remedy for Information Asymmetry”, Tom Fries, *Bertelsmann Foundation*;
- “Providing Independent Interpretations of Official Information”, Isao Matsunami, *Tokyo Shimbun*;
- “A Way to Save Time”, Pedro Markun, *Transparencia Hacker*;
- “Adapting to Changes in Our Information Environment”, Caesar Viana, *Goias University*;
- “A Way to Tell Richer Stories”, Sarah Slobin, *Wall Street Journal*;
- “An Essential Part of the Journalists’ Toolkit”, Cynthia O’Murchu, *Financial Times*;
- “A Way to See Things You Would Not Otherwise See,” Sheryl Phillips, *The Seattle Times* (Gray, Chambers, Bounegru, Ruetten 2012).

A number of researchers offer further definitions of data journalism:

Mirko Lorenz says it is a process that begins with analysis and continues with filtering and visualizing of data in a form that creates a connection to a story (Veglis, Bratsas 2017). According to Simon Rogers, data journalism adds tables, graphical analyzes of data and the biggest stories. It is basically the creation of news charts and includes elements of design and interactivity (Bradshaw, Lorenz and Rogers). According to Veglis and Bratsas (2017), data journalism is a process of extracting useful information from data, writing information-based materials, and building visualizations that help the audience understand the bulk of the story (Veglis, Bratsas 2017).

The process of data journalism

Different researchers describe the different stages of the creation of journalistic material through data journalism. According to Mark McCormick, describing the process for *The Guardian*, first comes data collection – that data may be officially published, sent to the media or collected from events that are currently happening. The first questions are what exactly data means, what to compare, and with what other data we can combine the existing one. Next is the filtering, part of which is unification of units, reformatting data that is in the wrong format, clearing unnecessary elements. After filtering, the results of the data analysis are outlined, and then, if necessary, it is double- or triple-checked. The final stage is publishing, and there are several options in this regard: visualization by programmers, use of free visualization tools, history writing, or simply – publishing the results (Data Journalism Handbook n.d.).

Paul Bradshaw recreates the reversed pyramid for writing journalistic texts through the prism of data journalism. According to him, the stages that the process goes through are the following: compiling, clearing, contextualizing, combining, and publishing (Bradshaw 2011).

Journalistic competencies

The new capabilities also require new competences and skills – in order to understand and produce certain results from the data, the journalist

should be a specialist in the field he/she is investigating. Here, technical skills are of great importance, starting with web-searching skills, downloading, reformatting, processing, retrieving information, organizing and visualizing. A great deal of the process also requires at least some beginners programming knowledge.

There are varied ways to visualize the results obtained from the data processing, whereas said ways are created using programming or ready-made tools. Visualizations may be represented by graphs, tables, interactive maps, interactive elements, small computer programs, specially designed for the purpose of the material in order to navigate through the data. Stella Konstantinova describes the impact of “Big Data” on the Internet, as well as the new professional roles for journalists, which include a graphic designer, a programmer and an interactive designer, connected with the processes of data journalism (Konstantinova 2015, 157). This provides a whole new palette of means by which journalists can present ideas and perform their professional mission. Iliana Pavlova describes the activity of the journalist-programmer, whose main function is data journalism (Pavlova 2016).

For the purposes of studying the necessary journalistic competences in the convergent media environment, the editors-in-chief at some of the major Bulgarian online media comment on the new trends related to data and data processing in deep-thought interviews:

Do you think it would be useful for your media if your employees were able to program? Can some of them program? How is it useful?

Programming and computer languages competence are related to the new trend in online media – practicing data journalism. To have quality results in data collection, analysis, and visualization, you need to have a high level of technical expertise, including being able to create small programs and having visualization software skills. According to respondents, this process is still barely noticeable in Bulgaria, and such competencies in the field are not mandatory, but everyone points out that these skills would be useful for employees in their media. Some are of the opinion that this is the future of journalism:

– In my view, this is a key competence and will become more and more important. We are trying to develop journalism with data. The issue of journalism with data is interesting. In my opinion, there are not enough

examples in Bulgaria because there is not so much data yet. This does not mean that we should stop working in this direction. Data increases more and more. At this stage, however, we are rather looking for collaboration between journalists and programmers, whereas the latter have to do the technical work after they have all achieved something together. In my opinion, this will be an important competence for journalists, Alexey Lazarov, Editor-in-chief at Capital. He answers the question whether he would prefer a candidate for a job, who speaks a foreign language, or if he would favor one that knows a computer language.

– There is a tendency to work with open data that is in theory out in the open but actually difficult to find. This is particularly true in public procurement and European projects. There's a lot of information published out there, but the problem is how to find it, see how to analyze it, check if this is the entire information or if there's something missing. Here, on the one hand, we have technical expertise because there are a lot of things to be read, and this is not done with the help of a simple search engine. There was a time when you could just look for something and it would come out, but now a lot of things are uploaded to PDF files, Vladimir Yonchev, Editor-in-chief at OFFNews.

– The next 22-year-olds I will hire, I do not know when that will be, but when they are here, they will need to know a computer language. This means being able to create a small page that shows the most interesting items in your article. That would be great. You can visualize data or make a database of all the people you write about. You can create very rich and interesting products as long as you possess this skill. There is a great joy in this – to create something on your own, Sibina Grigorova – owner of Webcafe EAD.¹

Technologies and tools

When collecting, organizing, and filtering data, a number of professionals in the area of data journalism recommend the program Microsoft Excel – popular precisely for its convenient and useful table and data handling capabilities. Another suitable tool is Open Refine, and for working with statistics – SPSS.

¹ Excerpt from Ivan Valchanov, dissertation thesis on: *Identification of the journalistic profession in convergent media environment*, 2017.

In terms of visualization, free and easy-to-use programs can be used to convert data into images or interactive elements or maps. The popular tools for visualization are Google Fusion Tables and Color Brewer, and Tableau and Data wrapper have also become very popular tools. In terms of mapping, the right products are Story Map JS, Google Maps, Mapbox and Open Street Map. For the more sophisticated visualizations and the creation of small computer programs that the audience uses, it is necessary to have computer language knowledge and the appropriate writing programs are TextMate, Vim and Sublime.

Conclusions

Although gaining momentum within a decade, large media brands are already looking at the new form of journalism – data journalism. These brands develop and complement the toolbox on a daily basis, creating new, diverse and useful media products for the audience.

Data journalism has its own features that distinguish it from the familiar “fast” online journalism. Because of data journalism, serious and in-depth investigations can be carried out, leading to new opportunities for investigative journalists and those who practice it.

The reviewed scientific publications on the topic, as well as the reviewed tools, methods and techniques in data journalism present data journalism development to date. Although not extensively researched and practiced in Bulgaria, this new form of journalism will become more and more important and will invariably find its place in the scientific and professional sphere. Future research on this topic could be focused on and develop each of the elements examined in this study and provide a more detailed analysis of data journalism.

References:

- ACEJMC Professional Values&Competencies (no date), <http://journalism.ku.edu/ace-jmc-professional-values-competencies>, accessed 30.01.2018.
- Bardoel J., Deuze M. (2011), *Network Journalism: Converging Competences of Media Professionals and Professionalism*, “Australian Journalism Review” 23 (2), pp. 91-103.

- Behind the Scenes at the Guardian Datablog*, in: *Data Journalism Handbook*, http://datajournalismhandbook.org/1.0/en/in_the_newsroom_3.html, accessed: 27.01.2018.
- Bradshaw P. (2011), *The Inverted Pyramid of Data Journalism*, <https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>, accessed 18.01.2018.
- Coddington M., *Clarifying Journalism's Quantitative Turn*, <http://www.tandfonline.com/loi/rdij20>, accessed 30.01.2018.
- Clark R., *The Pyramid of Journalism Competence: What Journalists Need to Know*, <http://www.poynter.org/2014/the-pyramid-of-journalism-competence-what-journalists-need-to-know/251048>, accessed 30.01.2018.
- Franklin B., *The Future of Journalism*, <http://dx.doi.org/10.1080/1461670X.2014.930254>, accessed 30.01.2018.
- Gray J., Bounegru L., Chambers L. (2012), *The Data Journalism Handbook*, O'Reilly Media.
- Gray J., Lucy Chambers, Bounegru L., Ruetten W. (eds.) (2012), *Data Journalism Handbook*, O'Reilly Media, Boston.
- Hewett J. (2016), *Learning to Teach Data Journalism: Innovation, Influence and Constraints*, SAGE Publications.
- Howard A. (2014), *The Art and Science of Data Driven Journalism*, <https://towcenter.org/wp-content/uploads/2014/05/Tow-Center-Data-Driven-Journalism.pdf>, accessed 30.01.2018.
- Konstantinova S. (2013), *Transformed Competence: Qualification Profile of the Online Journalist*, "UNWE Research Studies 2013" (in Bulgarian).
- Konstantinova S. (2015), *Online Media*. Avangard Prima, Sofia (in Bulgarian).
- Lewis S., *Journalism In An Era Of Big Data*, <http://dx.doi.org/10.1080/21670811.2014.976399>, accessed 30.01.2018.
- Mair J., Keeble R., Bradshaw P., Beleaga T. (eds.) (2013), *Data Journalism: Mapping the Future*. Abramis Academic Publishing.
- Newman N. (2016), *Journalism, Media, and Technology Trends and Predictions 2017*, <http://reutersinstitute.politics.ox.ac.uk/our-research/journalism-media-and-technology-trends-and-predictions-2017>, accessed 30.01.2018.
- Owen L. (2017), *Not a Revolution (yet): Data Journalism Hasn't Changed That Much in 4 Years, a New Paper Finds*, <http://www.niemanlab.org/2017/10/not-a-revolution-yet-data-journalism-hasnt-changed-that-much-in-4-years-a-new-paper-finds>, accessed 30.01.2018.
- Pavlova I. (2016), *Profession Journalist. 21st Century*, Faber, Veliko Tarnovo (in Bulgarian).
- Rogers S. *A Data Journalism Workflow*, <https://simonrogers.net/2013/01/27/a-data-journalism-workflow>, accessed 30.01.2018.

- Rogers S. (2012), *Data-Journalists Are the New Punks: Simon Rogers at TEDxPantheon-Sorbonne*, <https://www.youtube.com/watch?v=h2zbvmXskSE>, accessed 28.01.2018.
- Royal C., *At the Intersection of Journalism, Data Science, and Digital Media*, www.niemannlab.org, accessed 30.01.2018.
- Stampfl N., *Data Tells Stories*, <https://www.goethe.de/ins/bg/bg/kul/mag/20744960.html> (in Bulgarian), accessed 30.01.2018.
- Stray J., *A Full-Text Visualization of the Iraq War Logs*, <http://jonathanstray.com/a-full-text-visualization-of-the-iraq-war-logs>, accessed 30.01.2018.
- Veglis A., Bratsas Ch. (2017), *Towards a Taxonomy of Data Journalism*, “Journal of Media Critiques”, Vol 3, No 11, <http://mediacritiques.net/index.php/jmc/article/view/158>, accessed 30.01.2018.

Data Journalism Projects in Polish Mainstream Media Landscape. Case Study of Biqdata and Polityka Insight

Abstract: Data journalism has become recently one of the most important catchwords describing newest developments in journalism and media. Emergence of new (digital) communication technologies has created space for innovative, data-driven forms of journalism, seen as an alternative for traditional news-making. These innovations refer to almost all stages of journalistic processes, including collection, analysis, curation and publication of data. Dynamic development of data journalism, promoted by prestigious media organisations in the Western world (especially in Anglo-Saxon countries, but not only) also influenced journalistic practices in Poland. From 2013-2014 part of the biggest Polish mainstream media organisation started to introduce data journalism tools and develop data journalism projects. In this chapter emphasis is put on two of the most important initiatives including use of data journalism tools and practices, observed in Polish media landscape – *Polityka Insight*, developed by publisher of weekly opinion magazine *Polityka* and *Biqdata*, belonging to publisher of *Gazeta Wyborcza*, the biggest Polish quality newspaper. Although both projects are rooted in data journalism, their business models, target groups and general focus are different. In this chapter all these similarities and differences are described and analysed. Additionally, effectiveness of chosen business models is discussed, with special focus on paid content strategies and models, implemented by both analysed news media organisations – and seen as a response for visible crisis of traditional media and their institutional and economic foundations.

Keywords: Big Data, data journalism, innovations in journalism, legacy media, mainstream media, paid content

Introduction

Data journalism is frequently seen as one of the most popular concepts describing developments of journalism in past few years. Development of new technologies and digitalisation of media have enabled visible progress in terms of data harvesting, databases and infographics.

In the very beginning data journalism was generally seen as a domain of elite media outlets and leading global journalistic brands. Some data-driven stories produced by *The New York Times* ('Snowfall') or *The Guardian* ('Investigate your MP's expenses') became paradigmatic examples of this new journalistic practice. However, step by step this type of journalistic practices started to spread all over the world, and even some smaller news organisations started to experiment with this type of tools and activities. This way currently we can easily go beyond 'usual suspects' to research some interesting examples of data journalism practice. Since the early 2010s, this developments have been also observed in Poland, with some mainstream media organisations experimenting with new forms of news-making, based on raise of digital tools and practice of multimedia journalism.

Defining and understanding data journalism

Data journalism is generally understood as the journalistic collection, analysis, curation and publication of data (Coddington 2015). Origins of such phenomenon may be found in the 1960s, when Philip Meyer started to use computers to do investigative journalism and wrote his well-known book *Precision Journalism: A Reporter's Introduction to Social Science Methods*, first published in 1973 (Parasie, Dagir 2013).

Involvement of given media outlets in data journalism activities usually requires introduction of new specialties and specialist in newsroom's space and structure. Frequently traditional reporters are supported by specialists like graphic designers and statisticians. However, we can also observe the new types of journalists, combining traditional journalistic skills with some additional education paths (like mentioned graphic design and statistics) is emerging. In this respect additional question appears: is a specialist in research and data presentation still

a journalist? Such questions are asked by some experts and researchers of this field (Wawer 2016). In effect, purely journalistic character of some data journalism initiatives can be also questioned (Coddington 2015).

Considering media organisations involvement in data journalism activities, another big question appears, when we analyse economic context of such practices. As mentioned before, data journalism requires specific resources, which have to be collected and then applied by given newsroom. Usually it means substantial investment needs – in terms of money and human resources. On the other hand, data journalism activities are supposed to create added value for the users of media outlets services, prompting them to engage more with given media organisation and to transfer different kind of desired values (time, money) to them. This way, data journalism may be seen ‘as an opportunity for value creation through revised business processes as well as new products and services’ (Lewis, Westlund 2015).

Creation of such added value is closely related to ability of given media to offer products and services, which can potentially attract large number of users and/or fulfil some financially attractive niches on the given market. It means orientation towards mass audience or smaller but promising groups of users, like for example corporate clients.

In effect, different types of data journalism activities are held by various media outlets. Some of them use data journalism occasionally, others invest in much more institutionalised form of data-driven journalistic activities, like for example Datablog, offering regularly content produced by data journalists.

In effect, data-driven journalism include quite wide variety of practices and their institutional frameworks. Coddington (2015) provides detailed clarification of different journalistic activities related to data analysis, differentiating CAR (computer-assisted reporting), data journalism, and computational journalism. Uskali and Kuutti (2014) argued, that we can distinguish different streams of data journalism – investigative data journalism and general data journalism. They listed also three models for organising data journalism activities in the newsroom: ‘traditional data desk model’, ‘flexible data projects model’ and ‘entrepreneur model’ (called also ‘subcontractor model’).

Development of data journalism in Poland

Data journalism in Poland is much younger than in for example in Anglo-Saxon countries (Splendore et al. 2016). Polish media companies started to have interest in this type of journalistic activities in early 2010s, mostly following examples from other countries.

It's important to underline that data journalism wasn't introduced to Polish mainstream media environment by the biggest and most influential media organisations. First tries were made by economic and financial news website Forsal (belonging to Infor Biznes Group) in 2012, publishing data-driven infographics – in the beginning from time to time and regularly from fall 2013 (Małuch 2013). Their activities were appreciated by national journalistic community, as in 2013 Forsal was nominated to Grand Press Digital Awards, prestigious Polish journalistic award for innovators in digital journalism (Kuś 2016).

The biggest Polish media organisations started to implement some data journalism tools from 2013, when *Polityka Insight* was created by *Polityka* publishing group, as a sort of additional paid service, offered by publisher of opinion weekly *Polityka*. The biggest Polish newspaper *Gazeta Wyborcza* joined the competition in fall 2014, with creation of Biqdata, paid service with data journalism, led by Vadim Makarenko (Kuś 2016).

In Polish data journalism landscape we can also find examples of activities held by non-commercial institutions, from outside of institutionalised media field. For example, NGO Fundacja ePaństwo runs website mojePaństwo, providing data regarding different aspects of public life in Poland (Kuś 2016).

Gazeta Wyborcza and Polityka

– origins and forms of their engagement in data journalism

Daily newspaper *Gazeta Wyborcza* and weekly opinion newspaper *Polityka* are inherent part of Polish mainstream media environment. Both media outlets have belonged for the whole period after 1989 to the group of the most influential media organisations in Poland, in terms of opinion-making, popularity and influence on public opinion. Both media organisations achieved also quite stable financial position in that period. *Gazeta Wyborcza* is part of media holding Agora – and *Polityka* is published by company established by its own employees, being a sort

of exception in Polish media environment (Hofman 2014), dominated by big national media groups (public and private, like mentioned Agora) and big foreign investors, mostly from US and Germany.

More and more evident crisis of print media, which has been observed especially from the late 2000s, forced both media organisations to reinvent their business models and think about substantial investments in terms of building their 'digital capacities' in order to compete with other legacy media, migrating to digital platforms, and net-native media organisations, trying to compete versus 'old firms'.

In order to prepare for fierce competition in the digital era, both news organisations introduced their websites and promoted news types of journalistic activities amongst their employees, including blogging and active presence in social media (Kuś 2011). Later, after 2010, both organisations started to offer paid content online, following general market trends in Poland and globally and doing it in pursuit of lost financial stability.

Both media outlets and their managers were also interested in new developments in journalism, observed mostly in the Western world (mostly Anglo-Saxon countries). This interest was reflected in inviting foreign experts in digital journalism or participating in thematic events (conferences, seminars or workshops) (Splendore et al. 2016). It includes data journalism, which started to be widely discussed as an alternative for traditional reporting, especially from the early 2010s. In effect, both media organisations introduced forms of data journalism into their portfolio. However, different models and approaches were adopted in each case.

Polityka Insight

Polityka Insight offered by *Polityka* publishing house was the first project introduced by one of the most important Polish mainstream media outlets, including elements of data journalism.

The platform, seen as independent from *Polityka* weekly magazine and its online services, started to operate in April 2013. From the beginning, the project has been coordinated and led by Wawrzyniec Smoczyński, previously editor in *Polityka* magazine. Preparatory, conceptual phase started in early 2012 and main preparations, including formation of the team and content creation started in January 2013, according to Smoczyński. From January 2013 to April 2013, *Polityka* Insight team prepared around 500 analyses, which were available for first group of clients (Pallus 2013).

Taking into account traditional definitions of journalism and media we cannot call Polityka Insight purely journalistic initiative. Polityka Insight calls itself rather ‘center for policy analysis’ or ‘platform of knowledge’, covering mostly politics, economy and European affairs. We can discuss this difference between Polityka Insight and traditional journalistic forms on two levels.

Firstly, Polityka Insight is not designed for general public. It has been declared, that this services are aimed at a narrower audience of business, government and diplomatic professionals. It means mostly institutional (corporate, public and non-governmental) clients. Polityka Insight declares that their product is: *selective, precise and to the point. By choosing critical facts and figures we save our readers from information overload and deliver insights: causes, implications, relationships and interests. We project possible scenarios, solutions, risks and chances. Our findings are presented in a concise format, always on one page* (Polityka Insight).

Secondly, Polityka Insight staff is not dominated by journalists. From the very beginning around twenty employees form the team and most of them are analysts (mostly economists, lawyers and sociologists), usually with government or business experience. The team is divided into six sections: Politics, Macroeconomics, Europe, Business, Legal and Social Affairs (Polityka Insight).

In 2013, only three out of twenty employees in original team were journalists (Smoczyński and two experienced journalists working previously for *Gazeta Wyborcza*: Łukasz Lipiński i Wojciech Szacki) (Pallus 2013). Polityka Insight employees, according to their own declaration, are: *constantly following developments inside the Polish government, tracking economic indicators, trends in the real economy and pending decisions in Brussels. We are present at Sejm sessions, EU summits and key business conferences. Our analysts scrutinise think-tank output and process statistical and economic data* (Polityka Insight).

Biqdata

Gazeta Wyborcza was interested in data journalism from the beginning of 2010s. This interest was expressed in few different ways. Firstly, *Gazeta Wyborcza* started to experiment occasionally with some basic data journalism tools. Secondly, they also organised some educational events for their own employees, as well as for wider public (Kuś 2016). Thirdly,

one of the journalists, Vadim Makarenko, future data journalism project leader, spent some time in Reuters Institute for the Study of Journalism (at Oxford University) in 2013, as a journalist fellow, having opportunity to contact people interested and involved in data journalism activities, including journalists like Simon Rogers from The Guardian, founder of its Datablog (Kurasiński 2014).

In effect, in September 2014, Biqdata coordinated by Makarenko, started to operate (*Agora...*). Preparations were launched three months earlier. In the beginning, the team consisted of four employees: Vadim Makarenko, Piotr Kozłowski (journalist), Aleksander Deryło (graphic specialist) and Marcin Gębala (programmer and webmaster) (Kurasiński 2014). Later, number of employees has grown to six (Szews 2017).

As distinct from Polityka Insight, Biqdata always underlines journalistic character of the whole project and orientation towards general public. Contents produced by Biqdata team are available for all subscribers of *Gazeta Wyborcza* online version.

Makarenko's efforts related to development of data journalism in Biqdata were appreciated by journalistic community in Poland. In 2015, he was awarded with Grand Press Digital, the most prestigious Polish journalistic award for innovations in digital journalism (Biqdata 2015).

Two projects, two business models

Polityka Insight and Biqdata are both examples of paid content strategies, implemented by press publishers in Poland. This way publishers try to follow current trends towards development of paid content in print media industry.

However, these two projects represent two different business models and ways of monetisation of data-driven journalistic content. People behind both projects, for example Vadim Makarenko, also frequently underline existing differences between these two initiatives (Kurasiński 2014).

Polityka Insight

Launching of Polityka Insight has to be seen as a strategic move made by Polityka publishing house, in order to diversify sources of income for

the whole company and create completely new stream of their activities. Creation of such ‘center for policy analysis’ should be seen in this context.

We also have to underline significant investments made by Polityka publishing house in this respect. Creation of the team consisting of twenty employees and the whole process of launching of new platform are crucial elements of this strategy.

This substantial investments were made to create services, which might be attractive for significant number of institutional clients, mostly decision makers and executive-level professionals, expecting timely, regular analytical pieces of content, helping them to make daily decisions. Polityka Insight is also targeting users and clients, underlining three main groups of potential clients: corporations, public institutions and embassies.

Considering main services, called PI Premium, we can list four main types of products which are offered to clients: PI Analyses, PI Briefing, PI Knowledge and PI Publications. Additionally, apart from publications, Polityka Insight offers preparation of events for clients (Q & A sessions with decision-makers, expert seminars or open debates) (Polityka Insight).

PI Analyses include: *Five analyses a day on leading issues from the business and government agenda. Economic forecasts, sectoral trends, political developments and policy chatter from Brussels. Critical facts and figures, always in a one-pager format* (Polityka Insight).

PI Briefing is described as: *A daily headstart for those who need to be in the know. Our flagship product delivers crucial news from business and politics, a selective press digest, promotions at ministries and companies. A morning must-read for people on the go, delivered to your inbox* (Polityka Insight).

PI Knowledge refers to: *Critical knowledge whenever you need it. Over 4000 analyses, interlinked and searchable, aggregated around key topics. Constantly updated election polls and economic indicators. All in one place* (Polityka Insight).

PI Publications includes: *Timely briefs on complex topics of the day. Useful dossiers on mid-term policy issues. In-depth reports on systemic questions in politics, business and Europe. Delivered in PDF format, also to mobile devices* (Polityka Insight).

In sum, every weekday Polityka Insight subscribers receive a morning briefing and a set of five one-page analyses. Less frequently in-depth publications are released. Contents for each subscriber are personalised,

as PI offers four separate streams of data-driven reports for different target groups: managers (current economic forecasts, trends in key business sectors, profiles of emerging business players and digest of key consulting reports), diplomats (tracking of election polls and forecasts, daily political coverage, breakdown of party strategies and profiles of emerging political actors), public officials (in-depth analysis of government policies, international comparisons and best practices, digest of think-tank recommendations, profiles of key ministers and civil servants) and lawyers (tracking of legislation at ministries and parliament, early analysis of draft legislation and new bills, constant overview of changing regulations and profiles of public institutions) (Polityka Insight).

Polityka Insight services don't have fixed price. The price depends on individual agreements between Polityka publishing house and its clients. Clients can request two types of access: executive and institutional, 'available beyond a certain number of executive accounts, on the client's IP network'. Therefore, price of Polityka Insight services depends on two factors – scope of information needed by particular client and number of purchased subscriptions.

What's also important: Polityka Insight doesn't offer its contents for media outlets, in order to avoid uncontrolled dissemination of produced data-driven material.

Biqdata

In contrast to Polityka Insight, Biqdata may be seen as part of broader digital strategy of *Gazeta Wyborcza*, rather than separate division of the firm. However, we can also underline, that Biqdata was first ever 'digital only' service offered in the history of *Gazeta Wyborcza* (Szews 2017).

Considering pricing, Biqdata is not separated from other offers from *Gazeta Wyborcza* online, which uses metered paywall to monetise content produced by their media professionals. All the users of this online platform *wyborcza.pl* automatically get access to *Biqdata* articles, infographics and other materials. It's seen as part of the efforts towards creating added value for everyone paying for subscription.

In the beginning, Biqdata management planned to publish at least one story per day. Finally, number of published content items is lower (one story per 2-3 days) (Szews 2017), even taking into account fact, that Biqdata team grew from four to six employees.

Limited number of employees implies the fact, that most single stories are usually elaborated by the whole team.

Comparative analysis

Concerning all above mentioned variables we can easily find visible differences between both projects. Payments, pricing and distribution models of Polityka Insight and Biqdata are different. Two projects are also vary from each other concerning target audiences for their content. Comparative analysis of both projects is provided in Figure 1.

The most important difference between both projects are their target groups. Biqdata represents orientation towards mass audience, trying to attract the largest possible number of readers. Polityka Insight is designed for smaller group of users, mostly corporate clients and public institutions.

Figure 1: Comparative analysis of Polityka Insight and Biqdata

	Polityka Insight	Biqdata
Institutional design	PI operates as a part of Polityka publishing house, but is independent from the Polityka magazine. It has a separate team	Bigdata has a separate team, but it's part of general <i>Gazeta Wyborcza</i> online activities
Structure	Around 20 employees (mostly analysts, 3 journalists)	6 employees (3 journalists)
Target group	Narrower audience of business, government and diplomatic professionals (subscribers). No access for other media outlets	General public (subscribers of <i>Wyborcza.pl</i>)

	Polityka Insight	Biqdata
Publication strategy	Every weekday subscribers receive a morning briefing, a set of five one-page analyses and occasional in-depth publications	Every 2-3 days subscribers receive new piece of content
Pricing	Prices of subscriptions depend on type of institution and the number of prospective users	Fixed price (within Wyborcza.pl online subscription price list)
Paywall	Hard paywall	Metered paywall (within Wyborcza.pl online subscription policy)

Sources: *PolitykaInsight.pl*, *BIQdata.pl*.

Conclusions

Data journalism in Polish media seem to be still in early stage of development, especially compared to Anglo-Saxon media markets. However, we can also find some promising signals and symptoms, which can help data-driven journalism to flourish in Poland.

After the first year of activity of Polityka Insight CEO of Polityka publishing house Jerzy Baczyński declared, that all costs related to launching of Polityka Insight were returned, thanks to more than 100 institutional clients paying for their services.

First statistics of visits on Biqdata website were not very promising (less than 27,000 real users in October 2014) (*BIQdata.pl...*), but after more than three years its services are still offered to *Gazeta Wyborcza* subscribers – and it means that there is demand for this type of journalistic content.

Polityka Insight and Biqdata intend to be innovative not only in terms of using new tools and implementing new journalistic practices. Both

projects try also new paths in terms of funding. Paywall is basic monetisation strategy in both cases, but Polityka Insight and Biqdata tried also to experiment with competing for journalistic grants, for example within Digital News Initiative (Innovation Fund) supported by Google. Polityka publishing house got grant for development of their digital projects, including Polityka Insight, in 2017. Biqdata (in cooperation with team of *Gazeta Wyborcza* got grant for project “X-ray the politics” in 2016 (Press 2016; *E-Kiosk*...).

However, the fact that both projects try to rely also on grant money means, that there is still some uncertainty concerning economic and commercial potential of data journalism on Polish media market. The fact that apart from *Gazeta Wyborcza* and its Biqdata there is no other big mainstream media data journalism project dedicated to general public may be also seen as a confirmation of such doubts Polish media sector has in terms of potential profitability of such investments.

References:

- Agora uruchamia BIQdata.pl z dziennikarstwem danych* (2014), <http://www.virtualnemedia.pl/artykul/agora-uruchamia-biqdata-pl-z-dziennikarstwem-danych>, accessed 01.02.2018.
- Biqdata (2015), *Vadim Makarenko z BIQdata.pl z nagrodą Grand Press Digital*, <http://biqdata.wyborcza.pl/biqdata/7,159116,22114298,vadim-makarenko-z-biqdata-pl-z-nagroda-grand-press-digital.html>, accessed 01.02.2018.
- BIQdata.pl wystartował z 27 tys. czytelników, Agora zadowolona* (2015), <http://www.virtualnemedia.pl/artykul/biqdata-pl-wystartowal-z-27-tys-czytelnikow-agora-zadowolona>, accessed 01.02.2018.
- Coddington M. (2015), *Clarifying Journalism’s Quantitative Turn: A Typology for Evaluating Data Journalism, Computational Journalism, and Computer-Assisted Reporting, “Digital Journalism”,* 3(3), pp. 331-348.
- E-Kiosk, Fratria, Grupa Radiowa Agora i Polityka dostaną 345 tysięcy euro w czwartej edycji Google DNI* (2017), <http://www.virtualnemedia.pl/artykul/e-kiosk-fratria-grupa-radiowa-agora-i-polityka-dostana-345-tysiecy-euro-w-czwartej-edycji-google-dni>, accessed 01.02.2018.
- Hofman I. (2014), *Model redakcji zarządzania na przykładzie spółdzielni pracy “Polityka”, “Zeszyty Prasoznawcze”,* 3(219), pp. 472-492.

http://www.press.pl/tresc/42920,polityka_-agora_-natemat_pl-i-tok-fm-zdobyly-dofinansowanie-z-google-dni, accessed 01.02.2018.

Kurasiński A. (2014), *Vadim Makarenko (BIQdata.pl) – dziennikarstwo wyjaśniające czytelnikowi rzeczywistość jest przejściem naszego zawodu na nowy cyfrowy etap*, <http://blog.kurasinski.com/2014/09/vadim-makarenko-biqdata-pl-dziennikarstwo-wyjasniajace-czytelnikowi-rzeczywistosc-jest-przejsciem-naszego-zawodu-na-nowy-cyfrowy-etap/>, accessed 01.02.2018.

Kuś M. (2016), *Data Journalism in Poland – Practice and Education*, in: Szynol A., Konarska K. (eds.), *Media and Journalism in the Digital Era* (pp. 79-88), Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Kuś M. (2011), *Poland: Waiting For Citizens to Demand Online Media Accountability. Mediaact Working Paper No. 8/2011 on Media Accountability and Transparency*, Tampere: Journalism Research and Development Centre.

Lewis S.C., Westlund O. (2015), *Big Data and Journalism*, “Digital Journalism”, 3(3), pp. 447-466.

Małuch A. (2013), *Forsal.pl stawia na data journalism*, http://www.press.pl/newsy/technologie/pokaz/43668,Forsal_pl-stawia-na-data-journalism, accessed 01.02.2018.

Mejer J. (2016), *Polityka, Agora, NaTemat.pl i Tok FM zdobyły dofinansowanie z Google DNI*, “Press”.

Miller M. (2014), *Rozmowa z Vadimem Makarenką, twórcą i szefem BIQdata.pl*, <http://prasa.info/artykul/url,1155/>, accessed 23.11.2015.

Pallus P. (2013), *Rusza Polityka Insight – płatny serwis z analizami*, <http://www.wirtualnemedia.pl/artykul/rusza-polityka-insight-platny-serwis-z-analizami>, accessed 01.02.2018.

Parasie S., Dagir E. (2013), *Data-Driven Journalism and the Public Good: ‘Computer-Assisted-Reporters’ and ‘Programmer-Journalists’ in Chicago*, “New Media & Society”, 15 (6), pp. 853-871.

Polityka Insight, *Polityka Insight. Centre for Policy Analysis*, <http://www.politykainsight.pl/en/politykainsight>, accessed 01.02.2018.

Splendore S., Di Salvo P., Eberwein T., Groenhart H., Kuś M., Porlezza C. (2016), *Educational Strategies in Data Journalism – A Comparative Study in Six European Countries*, “Journalism”, 1(17), pp. 138-152.

Stępowolski R. (2014), *Jerzy Baczyński: „Polityka Insight” już osiągnęła rentowność*, <http://www.wirtualnemedia.pl/artykul/jerzy-baczynski-polityka-insight-juz-osiagnela-rentownosc>, accessed 01.02.2018.

Szews P. (2017), *Wizualizacja przekazu w dziennikarstwie. Infografika i data journalism w wybranych mediach*, <http://dspace.uni.lodz.pl/xmlui/handle/11089/23847>, accessed 01.02.2018.

Uskali T., Kuutti H. (2014), *Models and Streams of Data Journalism*, "Journal of Media Innovations", 2(1), pp. 77-88.

Wawer M. (2016), *"Data journalism": czy specjalista pozyskujący, porządkujący i prezentujący dane jest jeszcze dziennikarzem?*, in: Żyrek- Horodyska E., Hodalska M. (eds.), *Komunikowanie o nauce*, pp. 79-93, Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Crisis PR and Media: PR Tactics for Communication with Journalists and Media

„The secret of crisis management is not good vs. bad, it's preventing the bad from getting worse“

Andy Gibman

Abstract: The article is focused on PR tactics for working with media and journalists in crisis situations. It also aims to describe the challenges for PR experts to deal with crisis in the fake news era and to manage the crisis communication in social media. The text reviews Günther Bentelé's theory of public trust, and several studies from 2011-2017 about media trust, links between journalists and PR specialists, online crisis communication management, etc. Some questions are raised, related to whether we can speak about post-truth PR or about rumor management, how and when we can use humor in crisis communication, etc. In this paper are presented some case studies (PepsiCo, American Red Cross, etc.) for crisis communication management in social media. The last part of the paper is focused on the PR ethics in crisis by describing and analyzing the case of the BBC's famous presenter Jimmy Savile from 2012.

Keywords: crisis communication, crisis PR, media trust, PR ethics, fake news, rumor management

Introduction

The complex process of crisis communication management requires the corporations to include in their culture the approach of crisis prevention and risk assessment of the certain situations. This approach also implies a strategy for crisis communication management, which does

not rely solely on ready-made crisis plans. The knowledge-building process for the crisis requires the development of flexible crisis scenarios, where the PR crisis team needs to take under consideration the changes in the environment, the stakeholders and shareholders reactions and the media communication plan.

The crisis media planning can provide a specific content for the owned, paid, earned and social media. As one of the key target audiences in crisis communication, the media and journalists have an active role in the catalyzing or stabilization the crisis situation. The media trust issues in the last two years is sharpened because of the fake news easily spreading through the social media. On one hand, that can be considered as a prerequisite for the corporate crisis situation's emergence, but on the other hand poses new challenges for the PR experts in their journalists and media relations. In the presented examples in the text below, the fake news can be shaped by information taken out of context from a press release, an interview with exclusive manager, or a spokesman statement. The impact on the sensitive social media audience is almost immediate, and the company has to react quickly with a clear, factual and open statement on the case.

Knowledge-building process about the crisis

Crisis public relations field can be seen as an open social system, combining within the elements of prevention, rational response to emerging issues, strategic planning and programming. Managing the crisis situation requires a strategic approach and assessment that cannot be based on prepared nonflexible answers to the emerging organizational problems. Every crisis situation is unique, emerge and develop in different contexts and the PR specialists need to focus their efforts on the preventing phase, where they can control the problem not to become a long-term crisis for the organization. That approach excludes nonflexible crisis plans with ready-made construct answers and communication solutions.

The process of “knowledge-building” about the crisis includes the analysis of crisis situation that incorporates the elements: social environment analysis, analysis of organizational advantages, possible solutions and threats (SWOT or PEST analysis), definition of the actual problem

rooted in the crisis situation, and also its impact and implications on stakeholders and target audiences. Crisis communication management can be seen as process of building knowledge and know-how, because of the interdisciplinary approach, needed to resolve the crisis.

Even some of the crisis situation's definitions underline the importance of the "knowledge-building" theory in crisis situation management.

According to Aristotle, *the crisis develops itself mainly at a time of transformation – at the time when it's needed to be found a decision for the existing issues. The crisis narrative have past and future, causes and consequences...* (Pacheva 2009, 30).

The historian Randolph Starn thinks: *In the Greek tragedies, crisis situations were not only key moments in the process of change, but they became a moment of truth, the moment when you can see the true colors of the people or the events* (Starn 1975, 5, cited in Pacheva 2009, 30).

The crisis situation questions the legitimacy of the company, damages its reputation and undermines the target audience's trust. One of the main goals in the first crisis hour is to maintain trust in the organization's ability to cope with difficulties ahead. In that moment the important role has the established and preserved trust, because the preserved trust may be seen as intangible corporate capital to keep the company's legitimacy in society. An important role in this process is the work with the media and journalists during a crisis situation. Media relations are generally based on a two-way symmetric model of communication, which aims to achieve mutual understanding and awareness of mutual responsibility in crisis communication process. We need also to mark the fact there are tensions in the relationships triad: State Institutions-Organization-Media for the pursuit of the public opinion construction.

This "mediatization" of the crisis situation is characterized by the following:

- media coverage often overtakes corporate communication;
- the corporation is forced to choose the open or closed style of communication;
- the crisis has the elements of the news, it implies the re-telling of reality, offers a unique professional and economic opportunity for the media;

- journalists often find it difficult to decode the first messages coming from the companies about the crisis situation;
- on the other hand, the audience expects adapted materials and comprehensible content.

In public communication field, the following two aspects of crisis management are outlined: the organizations trying to reduce damage and negative effects on their reputation, the media tend to dramatize and overexpose the problem, and the cumulative effect leads to tensions in society and among the stakeholders. If the organization wants to “occupy the field of public communication”, it must take a proactive action, but also be aware for its social responsibility for resolving the crisis. When the company proactively spreads the corporate factual information about the crisis, it sets the tone, direction, and the pattern for future crisis resolution actions. That indicates the corporation is willing to be a source of information for the crisis situation and will openly co-operate with the media. In its first announcement, the company must take responsibility for the crisis, but without unnecessarily overexposing its own guilt.

There are five main principles (5Fs) for communication with the media also valid in crisis situations: fast, factual, fair, frank, friendly.

Volkswagen “Dieselgate”: Crisis PR strategy and tactics

The scandal with the Volkswagen AG called “Dieselgate” can be also illustrated with the application of the listed above principles. The management of the company wasn't very fast with the problem recognition, but they were fair in the communication with the media. The first statements of the company wasn't very factual, but mostly based on the company's main objective: regaining the costumer's trust and rebuilding the company's reputation. The factual information and the analysis of management's mistakes were published later on VW crisis website. The whole crisis communication strategy was based on the frank cooperation with media and journalists. The exclusive managers of the company (General manager and the exclusive manager in USA) at the time took the roles of the company's spokesmen, assuring the stakeholders that the company takes the issue seriously.

On 18 September 2015, the United States Environmental Protection Agency (EPA) issued a notice of violation of the Clean Air Act to German automaker Volkswagen Group. The accusation is that the company had intentionally programmed turbocharged direct injection (TDI) diesel engines to activate some emissions controls only during laboratory emissions testing.

What communication tactics the company decided to choose? On 22 September 2015, two days after the agency information, on the company's official YouTube channel was published the video statement of VW chief executive Prof. Dr. Martin Winterkorn (Winterkorn 2015). The company stops the selling of diesel cars in the US and pulls them out of the market. Martin Winterkorn takes the responsibility and resigns. He was replaced by Porsche's chief executive Matthias Müller who said: *My biggest and urgent task is to restore the people's trust in VW Group by doing what is necessary to solve the problems* (Müller 2015). The first company statement about the case highlights the principles of transparency for solving the problem, responsibility for the actions and also giving reference to the company's good reputation and trust. It also highlights the concern about the 600,000 employees who have been working with dedication to build the best cars for the customers and they shouldn't be blamed for the few people's wrong decisions.

Speaking in "one voice", communication consistency, keeping the friendly tone and the core message during communication in a crisis situation, are observed by the company. Michael Horn, who takes over as CEO of the company in the USA, reaffirms key messages in the company's official position. At the launching premiere of the new VW model Passat 2016, he raises the subject of the crisis and emphasizes that the company is taking responsibility and the board is committed to do what must be done to restore the customer's trust (Javier 2015). The management decided that the company will pay a voluntary \$ 1,000 compensation to every owner of a 2.0-liter TDI. They also launch a special fund managed by the lawyer Kenneth Feinberg (covering a financial compensation, free repairing and replacement of the cars, discounts and etc.). VW offers 3 years free tow and roadside assistance. VW and Audi lost their recognition for "Green Car of the Year", awarded from "Green Car Journal". For both brands, that means bad consequences for their reputation, because the diesel cars even received tax benefits. In the course of the investigation the company dismissed: Heinz-Jakob Noyer –

director of Volkswagen's brand management, Ulrich Hackenberg – director of research and development at Audi, Wolfgang Hatz – director of Porsche sports cars research and development and the PR expert Stefan Gruzem who "leaves the company by mutual agreement". On 27 September 2015, the company launched the special crisis website about the case in Volkswagen¹, offering news, live chat, questions and answers forms, phone line, tips for the car owners. On 10 December 2015, VW AG had the first open press conference. The journalists had to send their questions before the event, but also the media had the chance to ask their additional questions during the discussions.

Reputational and financial losses for the company in the early days of the crisis are serious, despite the proactive communication tactics. On the first working day, after the scandal (September 21st), the company's shares fell 20% on the Frankfurt Stock Exchange and the following days reached their lowest levels in 4 years. For the first time the company reached first quarter loss in 15 years. By September 27th, more than 34 lawsuits have been filed against Volkswagen in the US and Canada by VW and Audi car owners. The company is accused of violating the contracts, false advertising, violating federal and state laws by manipulating emissions data. Volkswagen AG has been removed from the stock indices Dow Jones Sustainability (DJSI – Sustainable Investment Index).

PR challenges in the fake news era. Crisis PR and new media

The cooperation between media, journalists and PR experts, when it comes to fulfilling their informative and persuasive functions, refers also to the issues about the trust in the media, in the organizations and in the whole system. Günther Bentele develops his theory for public trust, identifying 4 types of trust: one basic and 3 types of public trust. The basic trust is the interpersonal trust, and the public trust is divided into trust in personalities, trust in institutions and trust in the system. According to his theory, there is a particular role of the media in the process of building public trust. Media is more or less trustworthy as institutions, but also they can be considered as providers of trust in other

¹ WV AG temporary crisis site <https://www.vwdieselinfo.com/>, last visit March 2017.

institutions, organizations, personalities (Bentele 2000, cited in Zlateva 2008, 23).

Some of the research on the same topics – trust in the media, business, institutions, collaborations between journalism and the PR, registered similar decreasing trends in 2017. Edelman Trust Barometer 2017 pointed out that trust is in crisis around the world. There is a significant inversion. The general population's trust in all four key institutions – business, government, NGOs, and media – has declined broadly, a phenomenon not reported since Edelman began tracking trust among this segment in 2012. With the fall of trust, the majority of respondents now lack full belief that the overall system is working for them (Edelman Trust Barometer, 2017 Global Report).

Similar results have been reported by another study, which is more related to trust in the media. One of its aims is to examine the relationship between journalists and PR specialists. On 30 May 2017, in the Institute for Public Relations (IPR) website, were published the results from the Cision study (State of the Media Report) outlining the alarming tendencies for the trust decline in media (Ford May 30, 2017).

The journalists are concern how society perceives their professional field, because of the growing amount of fake news, disinformation and the need for facts verification in social media. A significant percentage (91%) feels like they lost public trust (The Cision 2017 State of the Media Report, 4). The journalists want news outlets and publications to be focus more on fact-based reporting than on the opinion-based coverage (The Cision 2017 State of the Media Report, 7). The PR experts should write angles that are less commercial and more usable for general audiences. Agencies and brands should produce in-house B-roll and sound clips relating to timely issues, especially visuals or videos of a work site or a specific product's use. If you are willing to write and distribute a press release you should be willing to reply to a few questions from journalists (The Cision 2017 State of the Media Report, 9).

There are already good practices and attempts to optimize relations between PR experts and journalist. The good example is the project of Ricky Yean and his agency Upbeat in San Francisco. They offer a digital alternative for journalists and PR experts' connections. The original email pitch allows to respond to whether or not the journalist is interested in the story directly from the email. When he/she hits "yes," it sends the per-

son to a webpage that includes the CEO's companies emails and phone numbers, so that he/she could get in touch directly, as well as all of the other info (press releases, past coverage, images, etc.). Most valuable PR resources for the journalists are the press releases, the stories, the images and the interviews. That allows tribal targeting, checking the sources and verification of information relevancy. Upbeat has built a database of several hundred thousand journalists (including freelancers and bloggers) and its software tracks what they write and tweet to see how their interests are changing (Owen 2017).

The researchers and PR expert's concerns about the media and public communications, are that the fake news are beginning to affect not only the political, economic and social spheres, but also the corporate sector. These concerns are not unreasonable, when PepsiCo had to quickly react to the misinterpretation words and very inaccurate citation of the company's CEO. The company's CEO Indra Nooyi was misquoted saying Trump supporters should "take their business elsewhere". It started from Conservative Treehouse, a blog that falsely quoted Nooyi's interview at The New York Times DealBook Conference (Liffreing, November 18, 2016). Not long after that, hashtag #BoycottPepsi, created by Trump supporters had over 19,000 mentions. Some media output the wrong quote in the headlines of their posts (Holmes, March 30, 2017).

What PR specialists can do in terms of optimizing their work with the media? There are a few principals that can be kept:

- assistance to the media that have been proven trustworthy – providing quality information so they can create a quality content;
- attention to the facts in their own press releases and especially to the events that rely on "storytelling";
- transparency for the researches and the financial support for them;
- developing a strategy for content marketing – better control over the message, reaching the target groups directly, creating authority, developing own corporate stories;
- not neglecting the small segments of target audiences;
- more opportunities for direct communication with the public and stakeholders.

In the post-truth era we can see examples for misquoting an information even from official press releases or even from the spokesperson's official statements. In January 2017, numerous news outlets reported a scientific opinion from May 2016 written by the European Food Safety Authority (EFSA). The results from the investigation about the toxicity of three specific classes of contaminants in food that are created by high-temperature processing of various vegetable oils, was cited in a press release form on 3rd May 2016. The media outlets don't mention the fact that only in high-temperature processing (over 200 degrees) the different vegetable oils, and the palm oil particularly, become toxic for the health. The most affected by this information was Ferrero, the company that produces the popular Nutella brand. So the spread by the media narrative "Nutella causes cancer", based only on this EFSA report, is untrue. In early January 2017, Ferrero announced that they would not be replacing the palm oil with something else, a key component of their product. The company has also launched the advertising video campaign for the product manufacturing process.

The cases above rise the question: Is it correct to speak about post-truth PR or we can say we need rumor management? Companies cannot and should not waste time, energy and resources to fight against every negative post that is published about them on the Internet. The people's hypersensitivity makes them choose to believe every message, if it promises to give them a ready and easy response to their problems. People will change their opinions very quickly if the promises do not offer solutions to their problems. Big social media channels do not want to risk their politics, their reputation and trying to adapt their algorithms and business practices to filter out the fake or misleading news. The rumor management process have the following algorithm: identification of rumor type, evaluation of the credibility, evaluation of the truthfulness.

That can be also applied for the social media rumor management. The social media plays a significant role in spreading rumors and the stories. That makes them a component for online crisis communication management. More than one-quarter of crisis spreads to international media within an hour and over two-third within 24 hours. It still takes average of 21 hours for companies to respond to the published information and after one year 53% of the companies had not seen share prices

regain pre-crisis levels (*Containing a Crisis. Dealing with Corporate Disasters in the Digital Age* 2013, 2).

The corporate communication's core, when it comes to online communication planning, is to be controllable or own a media channel: blog or corporate site. The benefits of that decision are two: the organization can take control over the information about the crisis situation and can also easily manage rumors and misleading facts about the crisis. The companies have one "golden hour" to identify the problem, to choose the strategy, to compose and launch the message. It is very difficult to follow this rule for suddenly occurring incidents, when representatives of so-called "citizen journalism" quickly publish photos from the events on different social platforms, before the company can react. In that situations, it is better to say a little than say nothing. In the first 15 minutes the following should be done: fact checking, organizing the emergency teams, writing the first message, preparing the statements from the management, keep thinking from the perspective of the media and the victims (what they need to know), and planning more detailed information.

In the first hour, social media messages from the companies can be based on less information, but need to show that action is being taken to resolving the problem. There are three *Cs* which are a good starting point for building the crisis message:

- *Care and Concern* for the victims and those affected by the crisis,
- *Control* of the situation,
- *Commitment* to resolving the problem (Griffin 2014, 216).

We can summarize four key issues about crisis communication in social media:

- Who is responsible for social media monitoring, who will collect and analyze the information received through it?
- Which social media users and followers' activities will be taking under consideration and will give them an answer?
- What advice we will give to our employees for using the social media? Whether they will be involved in communication as "company ambassadors", or their activities will be kept to the minimum level.

- Is there a need of particular employee's restrictive behavior during the crisis situation?

There are some tactical steps that company can follow in social media:

- Social media listening: follow the user's comments and active presence. When there is a problem, the simple publishing of a link with required information can prevent the escalation. Following the comments could point out the sensitive topics for the people and the company can be prepared better to respond in the future.
- If the company can't respond to users expectations, it's good to set the time frame for a future dialogue.
- The company must be open and honest: deleting the provocative and negative comments and reviews is not recommended, especially in critical situations. If there is a mistake made by a company employee, it should be admitted.
- Careful access to social networking accounts: just some people to have the access for administering pages and publishing information on behalf of the organization.
- Publishing instructions on the pages, regulating verbal behavior of users. They can clearly state what comments, images, and discussions will be removed by the administrators (Walter 11/12/2013).

Using humor in social media by crisis PR

The company can use a humor and informal approach in communication with the target audiences, when not facing a grave and tragic crisis situation. In 2012, the British telecommunication company O2 used the humor in its Twitter account for its crisis response after a 24-hour service disruption that left hundreds of thousands without a cellular connection. The crisis team handled the situation in a relatively small time frame. They even successfully redirected stakeholders seeking information to the company's Facebook and Twitter feeds.

Instead of the usual trivial corporate tone, O2 is trying to bring more humor and personality. An approach that would not be appropriate for all companies.

Customer (@24vend_Ltd): “@O2 had to travel to Italy to get signal -- desperate times!!!”

O2 response: “You can come back now. We’re back in business :)"

*Customer (@grahamcummings7): “@O2 F**K You! Suck d**k in hell”*

O2 response: “Maybe later, got tweets to send” (Lion, Georgiu 2012).

The company also has a spokesman who had to answer to all news media's questions, but the company uses actively its Twitter account to respond to every user comment, including those with inappropriate content.

The good example for turning the crisis situation into a positive social campaign is the Red Cross #gettngslizzerd Tweet in 2011.

Gloria Huang, a Red Cross social media specialist tweeted:

American Red Cross (@RedCross): Ryan found two more bottle packs of Dogfish Head's Midas Touch beer...when we drink we got it right #gettngslizzerd.

The tweet lasts an hour and was deleted, also Gloria Huang published an apology:

Gloria Huang (@riaglo): Rouge tweet from @RedCross due to my inability to use hootsuite...I wasn't actually #gettngslizzerd but just exited! #nowembarrassing (Wasserman 2011).

The American Red Cross explained the situation with the tweet in blog's company publication. Dogfish Head brewery was amused by the incident and took to their Twitter account to make a stream of remarks and retweets using the #gettngslizzerd hashtag, encouraging donations to the Red Cross: *Please join Dogfish Head Craft Brewery in raising money for the American Red Cross. If you're interested in donating a pint, please click here to learn more about Red Cross blood drives. Note: Alcohol can often make you more dehydrated. Dogfish Head recommends not drinking immediately before or after donating!*

The pubs and bars in 30 states serving Dogfish Head beer started the campaign “A Beer for Blood”. A tweet from a bar says:

Show us you donated a pint @redcross today & we'll buy you a pint of @dogfishbeer #gettngslizzerd (Preatorius 2011).

From the presented above cases we can make some conclusions:

- It is needed to quickly respond, because searching for a Twitter response is 15 minutes and about an hour on Facebook. The longer is the delay of the answer, the harder it becomes for the business to put out its point of view.
- Predicting the online crisis PR plan: companies are slow because they have been discussing their responses for a long time with the lawyer team.
- The excuse is an element from the crisis management plan. It can neutralize some of the criticism.

PR ethics in crisis situations

For many organizations, the question is not whether to engage in ethical communication during crises, but how and what factors influence an ethical stance, which influences crisis response. There is a need for a company to recognize the potential impact on the whole community during a crisis. Keeping the PR ethical code influences the communication and is related to the decision making process. The core values as trustworthiness, transparency and loyalty shape the communication strategy. Caring about the impact of the actions, respect for the shareholders and the stakeholders, fairness about the consequences that the company needs to take, are part of the building an ethical corporate culture.

The case with the BBC's long term journalist Jimmy Savile in 2012 is maybe one of the worst crisis situation, damaging the media image and reputation. The BBC was accused of sacrificing the journalistic standards to protect its own reputation after shelved an investigation against Sir Jimmy Savile – the BBC presenter, who had been accused of sexual abuse. The British police confirm that the already passed away TV presenter has committed more than 200 sexual abuses. Jimmy Savile, who died in 2011, was one of the BBC's stars in the 1970s and 1980s. For his charity, he received a knighthood from Queen Elizabeth II.

The BBC faced one of the biggest reputation challenges in their recent history. Newsnight's journalists reported on Twitter a few weeks later that another case of child sex abuse will be broadcasted on the program and after fuelled speculations it resulted in the wrong accusation against Lord Alistair McAlpine, a former adviser to Prime Minister Margaret Thatcher. The journalists made a new research and gave an official statement with apologies to lord McAlpine. On 2 October 2012, Newsnight editor Peter Rippon wrote a blog explaining why the BBC news programme had dropped its own investigation into Savile. The investigation by the police or by the Crown Prosecution Service had not established any "institutional failure" to make the Newsnight story weak. Savile is no longer alive to defend himself (Griffin 2014, 221).

The editorial decisions, and the way in which the communication was handled, led to the resignation of the Director General Gorge Entwistle during the news conference on 12 October 2012. The allegations of sexual abuse had followed the BBC star Jimmy Savile for years, but the problem got in deep for the BBC when they became aware that the commercial broadcaster ITV aired a programme called *The other side of Jimmy Savile* featuring interviews with alleged victims (Griffin 2014, 221).

After all, the BBC took the responsibility and tried to be transparent in the crisis communication. On October 22nd, the BBC program "Panorama" broadcasts *Jimmy Savile – what the BBC knew*. The show reviews and discusses Newsnight's collected facts and data for the crimes. There were launched two inquiries: the first one had to summarize if there were any failings by the BBC in how it managed the Newsnight investigation. It was led by the former head of Sky News Nick Pollard. The second had to describe and point out the violations of the BBC's journalistic culture code during the Savile's career there, and was led by the former High Court Judge Janet Smith. Newsnight editor had to step aside during the Pollard research.

The BBC wrote fact sheets and Q&A forms about the beginning and the development of the crisis, informing their audience about the gaps in their own work. Reports for the BBC's government issues are available on the website (Jimmy Savile report 2012) and can be downloaded. The reports accused the media for a "serious journalistic failure." The reports found that the fact checking wasn't correct, there was a lack of transparency and responsibility. The reports described

in detail the BBC News managers, editors and journalists and their replacements (Savile and Newsnight 2012). The info graphics include short biographical notes about their past positions, the responsibilities and the role they had in the crisis situation.

Winning trust as well as preserving it, is a long-term process. The BBC sacrifices its journalistic standards for trying to preserve its reputation. If we look at the crisis situation through the prism of the Günther Bentele's high and low values of trust, we will find that in this particular case the BBC has very low levels of trust: lack of competence, lack of communication, lack of public accountability and utilitarian ethical practices. They adopted positive responses after the crisis situation: competence to solve the problems through multiple changes in management, as well as communication transparency by giving citizens free access to the overall information by uploading everything to their website.

Conclusions

The 2017 Edelman Trust Barometer and Cision's State of Media Report marked the record declines of trust in media, institutions and authorities, alarming of the need to find a new approaches in B2B communication and rebuilding the relations with target audiences. Lack of trust reveals new challenges in the PR experts' work with the media, especially in crisis situation, when it is necessary for the communities and stakeholders to be informed correctly. State of Media Report outlined the relations between journalists and PR experts when it comes to fulfil their information and persuasion functions. After the fake news spreading in social media, the journalists are willing to give up form that to be first in the publishing of the news in favor of the quality and credibility articles. The media evaluates highly the PR materials with additional facts, video and story content, pictures, expert interviews.

The challenges for the PR experts in fake news era are related with rumor management and the assessment of the fake information impact on the company's reputation and audience's trust. The quick tactics for reactions in the first hour are the official statements and mobilizations of the communities, built after long term with brand community strategy.

The PR relations between experts and journalists during a crisis is based on the ethical principles, presented in the ethical codes of both professions. The principles of fair information, factuality, corporate responsibility, victim management, post-crisis communication are a part of building corporate crisis culture.

References:

- Bentele G. (2000), *Obshtestveno doverie v publichnata komunikatsia*, in: Zlateva M. (ed.), Public relations I konflikti v demokratichnoto obshtestvo (pp. 23-28), Sofia: Niks Print.
- Edelman Trust Barometer, 2017 Global Report, <http://www.edelman.com/global-results/>, accessed 31.01.2018.
- Ford E. (2017), *Journalism Credibility Crisis* | Cision 2017 State of the Media Report, May 30, <http://www.instituteforpr.org/journalism-credibility-crisis-cision-2017-state-media-report/>, accessed 30.01.2018.
- Freshfields survey (2013), *Containing a Crisis. Dealing with Corporate Disasters in the Digital Age*, <https://www.freshfields.com/globalassets/campaign-landing/cyber-security/containing-a-crisis.pdf>.
- Gatto G. (2016), *Nutella Ferrero Spot 2016 – Da 70 anni la qualità prima di tutto* [Video], retrieved October 24, 2016, <https://www.youtube.com/watch?v=8Jtu9gEreNY&feature=youtu.be>, accessed 30.01.2018.
- Griffin A. (2014), *Crisis, Issues and Reputation Management*, Kogan Page Limited, London.
- Holmes P. (2017), *Fake News is an Existential Threat to PR's Future*, March 30, [https://www.holmesreport.com/long-reads/article/fake-news-is-an-existential-threat-to-pr-s-future-\(part-1-of-2\)](https://www.holmesreport.com/long-reads/article/fake-news-is-an-existential-threat-to-pr-s-future-(part-1-of-2)), accessed 30.01.2018.
- Javier M. (2015), *Volkswagen's Apology on Tdi Scandal by President and CEO, Michael Horn* [Video], retrieved September 21, 2015 from <https://www.youtube.com/watch?v=dyiTwCuCRqg>, accessed 30.01.2018.
- Jimmy Savile – What the BBC Knew* (2012) [Video], retrieved October 22, 2012, from <https://www.bbc.co.uk/programmes/b01nspvr>, accessed 30.01.2018.
- Jimmy Savile Report: Full text* 2012, 12. 19, <https://www.bbc.com/news/uk-20782889>, accessed 31.01.2018.
- Liffreing I. (2016), *Fake News Sites Spark Pepsi Boycott, With Other Brands in Crosshairs*, November 18, <https://www.prweek.com/article/1416151/fake-news-sites-spark->

- pepsi-boycott-brands-crosshairs#B1OIFxF6DHpLBMFu.99, accessed 30.01.2018.
- Lion J., Georgiu A. (2012), *Calming a Twitstorm: O2's Masterclass in Dealing with 'Outrage Outrage'*, "WIRED" 17 July, <http://www.wired.co.uk/article/o2-outrage-social-media-masterclass>, accessed 30.01.2018.
- Matthias Müller Appointed CEO of the Volkswagen Group*, Press Release, 25 September 2015, https://www.volkswagen-media-services.com/en/detailpage/-/detail/Matthias-Mller-appointed-CEO-of-the-Volkswagen-Group/view/2726856/7a5bbec13158edd433c6630f5ac445da?p_p_auth=jKs7MN3I, accessed 30.01.2018.
- Owen L. (2017), *PR Pitches Are the Worst. This Agency is Trying to Make Them Better (And...I Like It)*, <http://www.niemanlab.org/2017/06/pr-pitches-are-the-worst-this-agency-is-trying-to-make-them-better-and-i-i-like-it/>, accessed 30.01.2018.
- Pacheva V. (2009), *Krizata pod control. Korporativna komunikatsia v krizisna situatsia*, Izd. M&M., Sofia: M&m.
- Preatorius D. (2011), *The Red Cross' Rogue Tweet: #getngslized On Dogfish Head's Midas Touch*, http://www.huffingtonpost.com/2011/02/16/red-cross-rogue-tweet_n_824114.html.
- Savile and Newsnight. Who Was in Charge* (2012), <https://www.bbc.com/news/uk-20297579>, accessed 31.01.2018.
- Walter E. (2013), *10 Tips For Reputation and Crisis Management In The Digital World*, 11/12/2013, <http://www.forbes.com/sites/ekaterinawalter/2013/11/12/10-tips-for-reputation-and-crisis-management-in-the-digital-world/>, accessed 30.01.2018.
- Wasserman T. (2011), *Red Cross Does PR Disaster Recovery on Rogue Tweet*, <https://mashable.com/2011/02/16/red-cross-tweet/?europe=true#Nzwm9U28zmqU>, accessed 30.01.2018.
- Winterkorn M. (2015), Video Statement, Prof. Dr. Martin Winterkorn, Volkswagen Group [Video], retrieved September 22, 2015, https://www.youtube.com/watch?v=wMPX98_H0ak, accessed 30.01.2018.
- Zlateva M. (2008), *Mostove kam konsensusa. Public Relations, jurnalistika i medii: problemi oblasti i konfliktni zoni*, Sofia: "Sv. Kliment Ohridski", Sofia.

Marketing Our Project Abroad

Abstract: The presented article can be treated as a guide, a bit as a scientific text, but first of all, as a textbook. It fulfills each of these functions in part and shows clearly that the reflection on the internationalization of business activities must cover various areas of social sciences and – among others – with particular emphasis on knowledge of marketing and its tools. The text distinguishes two marketing processes consecutive in the process of internationalization of a given company's business. The first is the preparatory stage, assuming work on the company's image, team building, preparation of tools and market perception. The second stage is the implementation of the project for international marketing activities, where the key role is played by widely described in the literature devoted to the marketing of tools. The most important part of the task is optimal preparation and properly selected tools. When the preparation phase is completed, a professional marketer starts the operation marketing and implementation of the whole project and timing is a matter of making a good decision: what does he want to start with, does he have the right team to complete this hard work, whether he starts at home or in a chosen place abroad; and he must learn quickly, both on good experiences and on mistakes.

Keywords: international market, promotion, company reputation, budget, website, direct marketing

Introduction

The fundamental part of the work that will determine all further activities when entering international markets is the preparation of a marketing plan. Every organization needs a marketing plan prepared especially for it, so the article only indicates general principles. The reader should treat them as a kind of list, which individual elements can be marked af-

ter implementation. If your organization deals with the tasks presented here, the result of all the hard preparatory work will be to put various elements into a fluid sequence of activities.

Preparation

Owner of the image

It is not enough to determine the image and reputation you want to describe on the international market, even for the internal use of the organization. As to the shape of this image, there must be a fairly broad agreement within the team working on the project. As a professional marketer who wants to achieve real impact on the course of things, you may have a tendency to positively redraw your image, but your fellow will quickly bring you down to earth. At the same time, your thing is to make sure that the organization's strengths are clearly shown. Any definitive introduction you make to your message should not resemble a definition, but rather express feelings or emotions. This should provide you with sufficient space in the future to build a strong image on the market, which – as if you do not look – is already a heavily overcrowded place.

If you do not demonstrate some unique and fascinating activities, you have little chance of success. For example, it is clear that the stand at international fairs or events should have the image you have planned. However, to achieve your professional goal, any other public manifestation related to your project – from a business card to a website – should emphasize this image and reputation. And in this very place, at least at big companies, things start to get complicated. What matters most to employees is the image of their own unit, department and boss (Nikodem ska-Wołowik 2004, 108). Many of them have a non-corporative website, produce their own non-corporate brochures, and some even push professionals from international marketing to recruit new customers on their own terms. As a result, we get by definition no results and a mess. In the worst cases, marketing changes into a scene between Laurel and Hardy, full of colleagues supporting both sides and showing lack of coordination and consistency. But in any case, this "fragmenta-

tion” of promotional activities will weaken the message, cause embarrassment in the market and, worst of all, seriously damage the image (Morreale 2015, 257).

How to deal with it? The stick and carrot method is usually the best combination with a large dose of patience. Ultimately, the metaphor stick – carrot (reward or punishment) turns out to be unfortunate in the group of friends working together on a project. A “flock of cats” could be a more proper term, so maybe we should use cream and sweet words, even in situations where completely different words go through your head.

- Take time for personal communication. Make sure you are in the right position to give your colleagues professional advice for free (Morreale 2015, 199).
- The ideal is to have a budget that allows you to finance units or other company publications (as long as they respect the image of the entire project, in other words – they listen to your advice), and whenever possible – to finance promotional publications throughout the institution.
- Make sure that each company unit is clearly visible (Ries, Ries 2003, 33).
- Explain why you are the editor of the whole, but do not join in endless discussions about translation or graphics (you employ professionals and that's their business).
- Ensure that you have real control over the content of the website, either directly or through cooperation with the webmaster convinced of your professionalism.

By accepting the above, you will find that a herd of cats, cream and sweet words will be much more effective than a big stick and sharp tongue.

Generally, it is advisable to have general rules, regulations and procedures to manage a sensitive area of promotion. Despite the best intentions, people are able to make serious damage to important and expensive activities at the institutional level through their ignorance (Kotler 2004, 10). Therefore, in special cases, it is advisable that the participants know that the stick is hidden somewhere here, but that you do not hold it, and someone else somewhere higher.

It is important that people get to you because they value your opinion and accept your advice. They should recognize your knowledge and experience. But using marketing jargon – a common strategy in such situations – can cause a reverse reaction (Morreale 2015, 112). Make sure you do not declare superiority – you have to deserve it.

Suitable promotional materials

Make sure all your materials and marketing activities are consistent with the image of the organization (Ries, Ries 2003, 31).

Once your organization has agreed on its image, use it to ensure that all of your marketing materials and operations are consistent with it.

- If you want your project to be recognized as young, innovative and dynamic, you will probably have to reach for a vibrant and contemporary design and youthful language (Ries, Ries 2003, 31).
- If your company is to be recognized as a place of art, your graphics will have to be sharpened with the style of communication that suits it.
- If, on the contrary, your company is old, respectable, known for high quality products, etc., and you want to attract seriously interested customers, your marketing materials must be more conventional.

Hire a professional graphic designer

Of course, it is not the role of a marketer from international affairs to create graphics, brochures and websites, and even – limit your ambitions in this matter and whenever possible – leave it to the professionals (van Rooijen 2008, 31).

But remember: it is up to you to clearly determine the content of the materials and give it to the graphic artist. This description must very accurately emphasize the image you want to design.

Make sure the graphic designer understands your idea well because some of them are overly creative and sometimes even tend to promote more of their own talent than your organization (especially when they start their career).

Importantly, it is also your role to ensure that the design you are working on with the graphic designer is consistent with all other company materials for the international market. All it takes is a few small “oddities” that a new website, leaflets or business cards will waste all their effort.

Slogan accompanying the logo (strap lines)

Most probably your venture uses a short slogan, which most often appears just below the logo. They are often a favorite marketing tool because they explain in a short way what an institution is doing or what it is.

The thought behind the slogan is an attempt to express the essence of the organization in short words (Ries, Ries 2003, 31). From a marketing point of view, you would like the slogan to bring the specificity of your firm to the market. Therefore, it should last for at least several decades, in some cases centuries. Even if this does not happen, it should arise with such awareness of its destiny.

The best slogans are very effective in image design. But although enterprises differ from one another, unfortunately, slogans tend to sound similar to each other, for example, “striving for excellence in communication”, or equally trivial. But you should not skip the company slogan, just because you do not like it. You have trouble if:

- No slogan: if your corporation does not use the slogan at all, it may indicate that it does not yet have a strong marketing tradition. Thus, it falls into the same category as other institutions without a specific style – today it is quite rare.
- The official slogan of your firm sounds like a dead language. Latin motto or ancient Greek may show you in the “antique” light, which could have a very interesting impact on the recipient, but you must be aware that today hardly anyone will be able to understand the slogan, and moreover that the image of the ancient company also has its drawbacks. Organizations that emphasize their history too much, can be perceived as haughty, elite, conservative, beyond contact with contemporary society (van Rooijen 2008, 34).
- The slogan is in your original language, and this is not a world language and is therefore unsuitable for the international market.

- The slogan is understandable only in the local context, because it has a special national character and – as such – will not function on the international market.

All these cases are a chance for you to show your creative talent (Lucian et al. 2008, 90):

- Start from scratch: return to your list of components of the venture image and reputation. Highlight on this list what you would like best to highlight when distinguishing your institution. This may be enough to generate new ideas (maybe not by you, especially by those who are more talented in this area, or who earn a living in this way).
- Check your ideas: try some of the results on a selected group of your existing customers: what do they think when they hear the proposed slogan? Do their associations correspond with your intention?
- Look for inspiration: of course you do not want to copy the slogan of another corporation, but you may need to know what others have done before. By the way – it is said: if you do not know how to steal a good idea, you should not work in marketing.

Additional slogans

As an addition to the main slogan – you (and/or the professional you hire) will look for other good slogans to use in advertisements, on postcards and brochures. Try to create a few:

- Write a few on the board in the office and ask others to criticize or add theirs, whoever they are.
- If you try to attract by being funny, be careful. Humor rarely works.
- If you try to shock, be even more careful. In another of your cultural contexts, it can take revenge dramatically (Ries, Ries 2003, 32).

When it comes to cultural sensitivity – your institution must be firmly convinced that you are the expert. It is extremely easy to commit a faux pas. The international market includes many cultures, each with a lot of taboos, so you can easily be misunderstood (Yalcin, Cimendag 2012, 9).

Although you have accumulated knowledge for years, always remember that you do not have to, and you should not even work alone. So, include others:

- If your company already works in foreign languages and cultures, ask employees of this department for help. They will be grateful if you ask them for advice.
- If you work in a specialist institution, you may not have direct access to such expertise, so you will have to work with colleagues, friends, interns or anyone with such specific cultural knowledge.
- Always verify your actions with others, remembering that sometimes such verification may surprise you.
- If, despite this, sometimes something goes wrong and you unknowingly offend someone, then at least you can prove that you did everything you could to avoid such an error (Morreale 2015, 290).

Watch out for language

If your company is located in a part of the world where English is not your first language, the good news is that it will be easier to pinpoint the difference between local and international communication as well as between local and international image and promotion. Your English-language brochures, websites, etc., will be exclusively addressed to international markets. However, if English is your native language, the same materials will have to handle all markets at the same time.

You might want to start using the language at a more advanced level, and thus, make the correction of your promotional materials more accurately if you target them to a specific international market. But you will only act if this market is large enough.

When addressing your materials to markets in their own language, with information of particular interest to this market, remember about the cultural specificity of these markets (e.g., using colors or images in materials for Ukraine or for China). However, some specialists disagree with this approach, claiming that in the future all markets should use English, so there is no need to take this extra effort.

Nevertheless, using a second or third language will allow you to prepare a tailor-made promotion and reach customers who may not speak a foreign language at the same level.

Of course, such a strategy requires employing professional translators and graphic overworking of materials. It will probably be very expensive to apply it on a large scale. But whatever you choose, do not save on employing unqualified translators.

Website

The company website has two main functions:

- provides information,
- stimulates “sales”.

It is quite common that some companies, acting in a public system, define themselves as the number one or “the most” in some aspect of their activity. From a marketing point of view, this is very unfortunate. Although many enterprises already have very good websites, most still suffer from information overload. What is worse, the website makes an impression as if it was overloaded with a combination of various firm printed information materials. And most often, this page has problems with itself, trying to achieve two different goals at the same time.

Also – especially at large ventures – because websites tend to arise and exist in this transient manner, the final result resembles a compilation mishmash. This is a serious problem that discourages potential customers (Marcos 2003, 9).

The company’s site must be well designed for international clients, but – more importantly – it must be built to stimulate the interest of future ones. So make sure that the people responsible for building and updating the website understand the requirements of the modern world (Ries, Ries 2003, 31).

If your site needs to cover the national and international market, because everything is in English (Spanish, French, etc.), it creates another topic – how to address the content and form of the website to both markets at the same time.

If you use English only in an international promotion, you should try to convince your webmaster to let you take care of its professional shape with international applicants, with the vision of international apprentices at the back of your head.

The website is also a good place to use the possibility of using

several languages. If, for example, China is one of your planned markets, you will want to build a website in Chinese. This is where your potential clients from this region will find dedicated information. You could ask Chinese students and graduates for Chinese-language comments that you will publish on this website. In any language, however, you will create pages, make sure that the translation is done professionally. As mentioned earlier, if pages in a foreign language are full of errors, your bold attempt will bring you more evil than good. Due to the fact that this is an electronic site, you should not lower your standards. On the other hand, as the reach of the Internet is much larger than the printed materials, you must be all the more careful and competent.

It is worth using the dynamic nature of the Internet. The potential of the Internet and its advantage lies in the possibility of frequent content updates. You can keep up to date on events in countries that will visit your marketing team in the near future or on visits to employees from your venture at local companies. “Meet us personally in your own country” or “Make an appointment” are very strong marketing tools. Use them fully when you are online (van Rooijen 2008, 40).

The Internet allows you to interact with your potential customers. Use the web professionally to deal with queries and then with potential clients. But be aware that you can become a victim of your own success. It seems very easy to invite you to e-mail ordering questions or ordering informers and folders. Start by building an automated response system that includes groups of standard explanations (FAQs). But what you really need is a system that will allow you to distinguish between really important questions of people who are seriously interested in your product and who are wasting your time with their questions.

Good product at a good price

The real challenge is balancing demand and supply. You must be exceptionally active in dialogue with international market (Belch, Belch 1995, 10).

If you already have the right image and promotional materials that express it, you are ready to go. However, first of all, make sure that your corporation offers a portfolio of products, specialties and services at an appropriate price level.

Prices driven by demand or supply? At most companies, producers define their portfolio. This means that the offered products reflect the experts and interests in the fields important for them. As a result, we receive a portfolio driven by supply, which may mean:

- In the best case, that the subject is at a high level of advancement, one step ahead of what is currently happening in a given field of science.
- In the worst case, that the portfolio is dominated by a group of empty, satisfied with their image of conditions, on average interested in clients' expectations.

Most international marketers are treated as supporters with the obligation to promote such a portfolio that the company generated. However, from a marketing point of view, it is much more attractive to have a portfolio built as a response to market expectations.

As a last resort, building a market-oriented portfolio is as unhealthy as a one hundred percent supply portfolio.

Demand is a variable and shaky thing. Corporations are not able to simply incorporate a new tape in place of the old one and smooth out all unprofitable products. In fact, both the human and physical structure of the company is not something that can and should be determined solely by demand. If the venture wants to enjoy a good market reputation, it should operate in various fields that are its strength, be able to develop new areas of offers or gradually correct old ones. Therefore, both views are correct and in the vast majority of cases the balance between supply and demand turns out to be a real challenge.

In international marketing there is a special need for high activity in maintaining dialogue with departments on new products, and constantly correcting the existing state of affairs so that they are more desirable for international clients (Belch, Belch 1995, 11). Your unit of development should show interest in receiving advice on the potential of the international market and it is in the interest of the company to formalize this process.

If you look at it, then – contrary to common practice – development units are certainly not the best place to set a price level.

Perhaps you believe that the final price is the result of costs, but this is only partly true. Knowing your own costs is certainly important. If

you do not want to finance international market, then prices should be above cost.

But of course costs can be calculated in many different ways:

- Full costs may include company expenditures on research or investments, raising prices above the reasonable level.
- Another extreme of valuation is the situation when having overloaded warehouses, services and exceeded numbers in offers, any income above basic costs brings serious additional profits. Although this approach seems justified, it does not seem to be a good basis for price calculations in the longer term.

The conclusion is that the right cost calculation does not have such secrets in it as your investor presents to you.

Much more important – than the above calculations – is the fact that the international market is a free market and the price level is primarily determined by market forces:

- First of all, you need to study how much other companies on the market demand for similar products or services.
- Secondly, you have to decide if your company (or particular products) is really positioning (in its higher, middle or lower part).

Some corporations are very effective in dumping prices. If it suits your market position to offer cheap average prices cheaper than your competitors, then such a pricing policy can be very effective. What you lose on a single payment will reward you with the number of customers.

Nevertheless, be aware that low prices suggest low quality. Such a strategy will be effective only if your venture is less focused on reputation and has the right size to deal with the crowd of not serious clients.

However, if the opposite is true, you should measure in the higher segment of the market. But be a realist: setting high prices and showing yourself as a high quality and elite place requires a really strong metric. What return can your company offer to an international customer on such an investment?

What is more, prices are justified not by the market position, but equally by the quality of services (availability of staff, standard of post-buy services, good communication). In all cases you will have to be able to demonstrate the value of the desired price: in usefulness

of your product, services, but also in for future contacts.

You must be a realist in assessing the value of branch of offered products on the market. You need to get involved in market research and in some cases in focus groups. Market intelligence plays a central role in pricing (Marcos 2003, 12).

It is not worth falling into the classic trap “if only we were cheaper, we would get a lot of clients”. It does not work like that. Someone may think that the international promotion of a new product, which has very low prices, is a dream for a marketer. But ultimately it turns out to be a marketing failure for many reasons.

Implementing

The result of your actions should be the same as in all other works: the most important part of the task is optimal preparation and properly selected tools. When the preparation phase is completed, start the operation marketing and implementation of the whole project is a matter of making a good decision: what do you want to start with, do you have the right team to complete this hard work, whether you start at home or in a chosen place abroad; and you must learn quickly, both on good experiences and on mistakes. If the idea of international marketing has been accepted and properly well-established in your company, the focus on activities will be moving more and more outside (Belch, Belch 1995, 12). If you do not have a managerial role in a large international marketing team, your tasks will probably be more operational than strategic. And if you or your company think that it will work, “cutting the corners” jump straight into the marketing plan without preparation, you can quickly see that you have limited your chances of success.

Market overview and situational analysis

Analyze your institution and the market where you want to act using SWOT techniques, that is, defining strengths, weaknesses, opportunities and threats. Strong points are inside your company. Opportunities and threats are external, and we usually analyze them using the PEST technique (politics, economics, sociology and technologies that create opportunities and threats on the market). These market analyses may arise using both preliminary data and deeper ones.

Defining a marketing strategy and establishing your own position

One of the marketer's main tasks is to find a common denominator for the recipient of our activities and our product or service. To achieve this, the specialist will use the STP technique (Segmentation, Targeting, Positioning) to identify and then interest the target group and position our product or service through the prism of seven Ps (Price, Product, Place, Promotion, People, Processes, Physical Evidence).

Defining goals

Each institution sets its individual goals, for example in internationalization – 20% increase in income in two years. On the other hand, the marketer's task is to set a more specific goal, detailing how this 20% will be achieved. Will the whole increase take place in the second year? In the first? In which countries? Based on these goals (usually determined on the basis of previous experience and knowledge about it, where external opportunities and internal strengths exist) the marketing strategy will be determined (Belch, Belch 1995, 8).

Direct marketing

Marketer uses his contacts with potential customers to build a relationship and try to sell the product. If you do it correctly, direct marketing can be extremely effective. It's worth, though, to remember that you must make the right offer to the right (Belch, Belch 1995, 14) person at the right time. To make sure you do not send out spam, offer an easy way to communicate feedback. Participation in branch fairs is sometimes seen as direct marketing and sometimes as a sales promotion.

Advertising

A paid promotion of goods, services, companies and ideas by the sponsor disclosed in the advertisement. Marketers place ads in all media since Internet TV. In the international marketing of any product most advertisements are purchased in branch specialized journals or special interest websites (Marcos 2003, 8).

Public relations

Planned and sustained attempt to establish contacts and gaining goodwill and mutual understanding between the organization and the sur-

rounding groups (Kotler 2004, 6). PR starts with good relations with the media. If there is something good and interesting to say about your company, its new products and achievements, make sure that professionally written press release is available and widely distributed. Other PR tools, such as sponsorship or organization of events, will also positively present the marketing image of your venture. PR requires more creativity than advertising, but it can be implemented at significantly lower costs and it brings generally better results.

Promotional sales

These are a short-term incentives that motivate consumers or distribution channels for the immediate purchase of goods or services. Usually promotional sale offers something special for a lower price for a short time to force the consumer to make a purchase decision faster than he/she usually does (Ries, Ries 2003, 30). You can also further motivate recruitment agents to work harder for you than for other clients by offering the periodically increased commission.

The correct and effective use of these four tools, however, requires a good CRM (Customer Relationship Management) application or an efficient management system as well as a dynamic website. Integrating your activity with all these tools will bring you better synergy and a higher “copy in euro”.

Operational plan and schedule of activities

After tuning the marketing mix with communication activities – according to your best options – it's time to create an operational plan for the nearest, usually three to five, years. Which of the activities will you take in next year? Why? Will these activities lead to the goals? What resources (personal, technical, financial) will you need? The operational plan is your guide – determines what you will do, when and why (Kotler 2004, 8).

Budget

Based on the business plan, we are able to determine the budget for approval by the company management. Make sure it contains all additional items regarding sub-rented external companies as well as information related to commissions that you will have to pay abroad marketing agencies or your representatives. For purely accounting reasons, it

is recommended to distribute the budget not only for communication tools marketing and other required costs, but also for geographical regions (van Rooijen 2008, 58).

It will help you decide whether to buy advertisements or participate in fairs in some countries and regions during your daily work, and you will have a good indication of whether your budget can withstand the burdens associated with this part of the world.

Conclusions

All marketers accept certain assumptions and have specific expectations, planning your activities. However, it should be borne in mind that our assumptions can be erroneous, and the development of the situation and the results of our activities are often different from expectations. It is therefore particularly important to predict which sensitive facts or situations can have a significant impact on the outcome (Kotler 2004, 2). As an example, let us use the periodic implementation of the online application. Having such a solution causes that the application process will be more efficient and your team will be able to work on other marketing tasks. However, if such an implementation is moved in time, then your other activities will also be delayed, the team will work in permanent short-term, and all of this will affect its goals accepted in the marketing plan.

The adoption and application of the suggestions and solutions presented above may – though not necessarily – significantly facilitate the implementation of such a serious task as the emergence of new markets. However, because the number of variables, the appearance of new elements and the dynamics of economic life may exceed our expectations, let us also be ready for the emergence of crisis situations. Their identification is important for two reasons:

- if situations or facts are getting out of control, you have the chance to intervene when they appear;
- if things are out of your control, you have time to warn your management that it will not be possible to achieve the goal.

In addition to identifying doubtful situations, it is also important to create a plan for unpredictable and critical situations that can always happen during the course of a lifetime.

References:

- Belch G.E., Belch M.A. (1995), *Introduction to Advertising and Promotion. An Integrated Marketing Communications Perspective*, Chicago: R.D. Irwin, Inc.
- De Wit B., Meyer R. (1994), *Strategy: Process, Content, Context. An International Perspective*, Minneapolis and St Paul: West Publishing Company.
- Kotler P. (2004), *Ten Deadly Marketing Sins. Signs and Solutions*, <http://media.imaxws.net/912/10deadlymarketingsins.pdf>, accessed 29.12.2017.
- Lucian R., Barbosa G., Filho J., Pereira F., da Silva I. (2008), *What Do Strategists Have in Their Minds? The Use of Structural Equation Modeling to Understand the Strategy Process*, Brasilian Business Review.
- Marcos A.M. (no date), *The e-Marketing in the New Economy*, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.476.101&rep=rep1&type=pdf>, accessed 30.12.2017.
- Morreale S., Spitzberg B., Barge J.K. (2015), *Komunikacja między ludźmi*, Warszawa: PWN.
- Nikodemcka-Wołowiak A., Górska T., Wołowiak M. (2004), *Nie tylko logotyp. Wyróżnienie i przynależność w biznesie*, Bydgoszcz, Gdańsk: Oficyna Wydawnicza Branta.
- Ries A., Ries L. (2003), *The 22 Immutable Laws of Branding*, London: Profile Books Ltd.
- Rossiter J.R., Percy L. (1997), *Advertising Communications and Promotions Management*, Boston: McGraw Hill Irwin.
- Siwińska B., Zimnak M. (2016), *Czas internacjonalizacji. Wyzwania dla polskich uczelni*, Warszawa: Elipsa.
- Stern L.W., El-Ansary A.I. (1992), *Marketing Channels*, Englewood Cliffs: Prentice Hall.
- Thompson A.A. jr., Strickland III A.J. (1995), *Crafting and Implementing Strategy*, Chicago: R.D. Irwin, Inc.
- Usunier J.C. (1993), *International Marketing: A Cultural Approach*, London: Prentice Hall.
- van Rooijen M. (2008), *Marketing Your Institution Internationally*, Amsterdam: EAIE Professional Development Series for International Educators.
- Yalcin E., Cimendag I. (2012), *Global Marketing Advertising with Cultural Differences: How can global companies Better Address Cultural Differences in Marketing Advertising*, Jönköping: Jönköping International Business School.
- Zimnak M. (2015), *Promocja edukacji akademickiej w Polsce po roku 2000*, Kraków: Universitas.

Internet Astroturfing and Security Threats: The Protests in Defence of Poland's Courts, Bots and Disinformation

Abstract: This article deals with the chaos that marked public information about the protests in defence of Poland's courts in July, 2017. These events have been interpreted as a case of astroturfing, both in the real and virtual worlds. An astroturfing campaign is an unethical marketing procedure, aimed at creating the impression of a grassroots movement, when in fact it is a co-ordinated campaign. Those at both ends of Poland's polarized political spectrum have called out their rivals on this kind of activity. This article is both a case study and an attempt to answer the question of whether, in Poland in 2017, we were the witnesses and targets of astroturfing. The question arises as to whether this type of activity indeed poses a security threat. This paper is an English translation of the author's article in "International Security Yearbook 2018".

Keywords: politics, astroturfing, digital marketing, manipulation, protest

Introduction

Michael Delli Carpini and Scott Keeter assert that better-informed citizens are better members of society, at least in terms of the standards of democratic theory. Participating more effectively and consciously in political events, they are able to relate political decisions to their results – depending, however, on how broadly they have been informed (1996) Henry Jenkins (2006) writes of “participatory culture”, a distinct ele-

ment of the Web 2.0 generation characterized by the harnessing of the potential of collective intelligence through the sharing of the social and intellectual capital of members of networks. John Zaller (2003) writes expressively of the “monitoring citizen”, emphasizing this person’s controlling role in the process of communicating about politics. The question may be asked, however, as to what extent these changes have really influenced receivers of information, or the ability to select and verify it. It seems that it is very difficult to manipulate today’s Internet users, highly skilled in digital matters and thus able to identify such attempts (Hille 2014). Technological changes are shaping a new type of citizen: the netizen (net + citizen). The active user introduces content to the web, sometimes only for fun, but sometimes also to comment on a given situation (Jenkins 2006, 72).

The great challenge facing today’s media is to serve an individualized, post-industrial society. We may observe the slow transformation from channels into platforms with content collected from various sources. These platforms work, or will work, in a personalized manner in a non-linear, interactive Internet environment. Of course, changes in areas such as the reception and distribution of media are taking place in stages, but fairly quickly all the same. From these changes arise both positive phenomena, and certain threats. On the other hand, paradoxically, the ease of obtaining information may in some way discourage recipients from looking for it. From yet another perspective, if the choice of content depends entirely on the recipient, then those who are not interested in political topics can very effectively avoid them. Then again, it can be argued, that at one point they could have done so, but for the very want of other options, they were “forced” to take in content related to politics.

The *diktat* of the receiver has changed this state of affairs (Prior 2007). The great popularity of social media and blogs has allowed citizen-journalism and alternative journalism to flourish. This has provided for increased variety in sources of information, but also undermined public trust in the media. Alternative media with a strong populist tone have gained popularity, being opposed to political and social elites and interpreting reality with a focus on alleged conspiracies. One of the effects of the growth in the popularity of new forms of communication is the possibility of conducting an astroturfing campaign on the web. The article presents a case study on the use of the term “astroturfing” during

the protests in defence of the Polish courts in July 2017. The objective of astroturfing activity is to create confusion and chaos. It is a fairly inexpensive undertaking, quite effective in view of the extent of its impact. It is very difficult to identify the sender responsible for such activity, and so to prove that such a campaign exists. As a result, such campaigns can be used in the case of asymmetric attacks. It is highly probable that they took place in Poland in July 2017, above all on Twitter but also in other social media, such as Facebook. These activities were related to our country's polarised politics. After characterising the narrative conducted in the media on astroturfing, it is possible to identify distinctive elements of tabloidisation (Piontek 2011), which serves the purposes of astroturfing by increasing social division and chaos.

The tabloid message is based on strong contrast, and its use in a sharply divided society and political class can bring measurable results. The use of the possibilities of new technology is now a common practice in the war of information. There are many indications that an astroturfing campaign has taken place in Poland. The question arises, however: who inspired it? Was it George Soros, at whom politicians in the government camp have pointed their fingers? Or, as several journalists and researchers have suggested, was it Russia? (Kovic 2016). A further interesting issue is the effects of this campaign: chaos, noise and disinformation (Klotz 2007, 14). All of these elements of the case study will be dealt with in the following text.

The work's theoretical basis, as well as examples of astroturfing campaigns, will be presented. The author will also show the context and narrative related to the presented case in leading online media, within a given time frame (20.07.17-30.07.17). The research part presents three important reports from various centres of Internet data analysis, as well as data that the author obtained by using the Google Trends tool. All of the indicated elements lend themselves to a deeper analysis of the occurrence of Internet astroturfing in Poland, with relation to the described campaign.

Astroturfing: definition and cases

Nowadays people exist in an unusually unstable, changeable hybrid world, where making forecasts is difficult. Technological development

is proceeding very quickly. New technologies give us the impression of relationships with other people – with users – but in fact individuals operate in an atomised environment and face the problem of loneliness. Forced in a way to absorb a great deal of information, we feel a sense of glut, or information overload.

Astroturfing has been well-known for a long time as a way for individuals, political parties and interest groups to conduct campaigns (McNutt 2007, 172-173). The basis of astroturfing activities is giving their audience the impression that the activities are part of a spontaneous, grass-roots movement. What lies at the basis of such activity is fraud and manipulation. Typical of astroturfing are business, political, social or religious campaigns that are co-ordinated, planned and financed by various interest groups. This activity is presented to the public as spontaneous, uncoordinated and “bottom-up”. Behind such false civic or consumer movements are often interest groups that employ professional public relations agencies. It can be concluded, then, that astroturfing can be called an artificial grass-roots movement. The Internet thus provides a very convenient space for such activity.

One of the better-known cases of business astroturfing was the activity of the American store chain Wal-Mart in 2006 (Peeples, 27). The company hired the Edelman PR agency to conduct a campaign that built a positive image of the firm, discrediting negative opinions. The group Working Families for Wal-Mart was created, with its members recruited and paid by the Edelman agency (scmagazine). The group presented itself on the web as a grass-roots Internet community, drawing together local community leaders. One of its leaders was the former mayor of Atlanta, Andrew Young. The idea was to create a grassroots group of influential members. An example of a person employed by Wal-Mart was the blogger Laura St. Claire, who travelled with her partner throughout the United States in a camper and reported her visits to Wal-Mart stores in a blog, “Wal-Marting Across America” (Bowen 2014, 219). Only later did “Business Week” journalists uncover the truth about who was financing the blog. Its content had been inspired by a PR agency (TVN 24).

Astroturfing is neither spam, nor the buying of so-called “likes” on social networks, nor trolling (Chu 2014, 21-22). Its effects are similar to such activities, but the condition of Internet astroturfing is the use

of bots. Those who “sow artificial grass” normally do so as a group, as such activity is very artificial, or through the use of bots. A bot is a program or application which, properly programmed, replaces a person. Bots may serve Internet customers or conduct our e-mail correspondence, but they can also be used to spread spam and viruses (Newman 2017, 15-17).

Astroturfing campaigns make use of the Internet’s ability to host viral activity. This is well illustrated by the actions of NATO forces in Afghanistan. In September 2011, after Taliban forces staged attacks near a NATO base and the American embassy in Kabul, the NATO command decided to step up its Internet activity, especially on Twitter (Bernatis, 2014). A false Afghan Taliban account was created, as well as many personal accounts, ostensibly of Afghan citizens. A discussion began about the official International Security Assistance Force – ISAFmedia@ account, in which bots took part as well as people (Sweetser 2014, 209). Special computer programs and applications were used, creating characteristic artificial accounts – bots.

Fake Twitter accounts are known as “sock puppets” (Bu 2013, 367). They can be made to publish automatic content, manage correspondence if the author is not present, and serve customers on the web. They can also spread spam and viruses, and conduct cyberattacks. Bots operate in social media, most often on Twitter (Howard 2016). The American example has shown that social media can be a weapon and that information warfare, while effective, also has an ethically debateable dimension (Sweetser, 209).

#3xveto, stop #AstroTurfing, #wolnesądy : media with various political affiliations

After Parliamentary and presidential elections in Poland in 2015, the Law and Justice (PiS) party won a majority of seats in the Polish Parliament, as well as the President’s office. PiS is viewed as conservative and populist. The new President, who should be impartial according to the provisions of the Constitution, has openly acknowledged his ties to the governing party. After winning its majority, PiS almost immediately began the reform of public institutions, reflected in the bills of law concerning the judiciary. Poland has split into two polarizing camps: the government’s, and the opposition’s. Those acting against the governing power’s measures are not only Parliamentary fractions – the Citizens’ Platform (PO), Modern (Nowoczesna) and the Polish Peasants’ Party

(PSL), but also extra-governmental organizations such as the “Committee for the Defence of Democracy” (KOD) and “Citizens of the Republic of Poland” (Obywatele RP) (Gromadzki, 2018).

The Polish case under study involved demonstrations following three bills of law that were approved by the national legislature on July 12, 2017, concerning Poland’s courts. These proposed changes to the judicial system set off a wide social response, which peaked in street protests in various cities, not only the largest ones, where demonstrations and so-called “chains of light” were organized. As a result of these protests and consultations with the legal community, President Andrzej Duda decided on July 24, 2017 to veto two of the bills, concerning the Supreme Court and the National Judicial Council. He signed the third bill.

In the middle of July, when protests began in the media, the term “astroturfing” appeared. Articles in pro-government media, especially between July 20 and 30, suggested and even demonstrated that the protests did not have the character of a grass-roots movement, but were controlled. On the other hand, articles in other media, not associated with the government camp, showed that in fact the Twitter campaign was online astroturfing, and not a demonstration in defense of the courts, known on Twitter by the hashtags #3XVETO and #WolneSądy (“#Free courts”). The Twitter campaigns #AstroTurfing, #StopAstroTurfing and #StopNGOSoros were meant to unmask the demonstrations in defense of the courts as an act of manipulation.

Below are extracts from the texts of articles and other online postings, interpreting the events on the web and in the streets in different ways. They appeared on Internet media between July 20 and August 30, 2017. The extracts were selected on the basis of publications contained in the four largest horizontal portals: Onet.pl, wp.pl, Interia.pl and gazeta.pl (WP, CBOS 2017), the websites of the most popular television stations, TVN 24 and TVP Info, (CBOS 2017) the government-affiliated stations Radio Maryja and TV Republika, (CBOS 2017) the three most important newspapers, “Fakt”, “Super Express” and “Gazeta Wyborcza”, and the right-wing web pages Nieuzałżna.pl and wPolityce.pl (WP, CBOS 2017).

Table 1.

wpolityce, niezależna.pl, TVP, Radio Maryja, TV Republika (associated with the government camp)	Wp.pl, Onet.pl, TVN24, Fakt, natemat.pl, (rather not associated with the government camp)
<p><i>These are not spontaneous protests, at the homes of PiS politicians; constant provocation of (PiS leader) Jarosław Kaczyński, meant to lead to emotional disorder and the break-up of PiS into factions. These are only some of the steps announced by the head of the Council of the Open Dialogue Foundation. Bartosz Kramek has devised them based on the experience of the Ukrainian Maidan movement (wpolityce.pl).</i></p> <p><i>Astroturfing. A very dangerous foe has entered Poland, and has not appeared by accident (niezależna.pl).</i></p> <p><i>Observing the exemplary organisation of the latest activity at the Sejm and the presidential palace, the communicative coherence and perfect functioning of the media as well as Internet trolls, it is difficult to avoid the impression that a staff of experienced and professional advisers stands behind this industry of manipulation (wpolityce.pl).</i></p> <p><i>Poland, today, is the object of massive disinformation on the part of the opposition forces,' said Jacek Sasin.</i></p> <p><i>Andrzej Dera: 'it's obvious that this activity is controlled. One has to wonder who is inspiring these attacks on the Internet' (tvp info).</i></p> <p><i>One of the television stations has taken the initiative of broadcasting live the events on the streets of Warsaw. That is no accident: you need unprecedented resources and money for this. There stands behind it a powerful milieu of lobbyists, to whom it is distasteful that we have cut the petrol mafia down to size, and reduced the</i></p>	<p><i>The PolitykaSieci.pl website has drawn attention to the #AstroTurfing hashtag that is spreading quickly in the web. The (astroturfing) concept signifies organized activities, which viewed from outside might seem spontaneous. The editors of the portal have associated the term with commentators, appearing in great numbers on the Polish Internet, who are not Polish users. According to this website, posts with identical content, critical of the Polish authorities, are appearing on TVN24's Facebook profile. According to the findings of the website, none of these commentators' accounts are related to Polish profiles, nor to networks of friends (WP).</i></p> <p><i>Day after day for nearly a week, thousands of Poles have protested massively in defence of independent courts. Peaceful demonstrations have been organized not only in large cities, such as Warsaw, Cracow and Poznań, but also in many other, smaller centres. Their participants have one demand: the veto of the PiS bill concerning the court system. Still, PiS supporters argue that the protests of recent days are not spontaneous.</i></p> <p><i>To explain and rationalise them, they have reached for that mysterious and menacing word: astroturfing. Among others, Anita Gargas of TVP (formerly of TV Republika) has used the term in the context of the demonstrations in Poland (gazeta.pl).</i></p> <p><i>Astroturfing? It is certain that both sides have been fighting. But the Internet is not going to change who is in power. Not yet (wyborcza.pl).</i></p>

crime related to the collection of VAT tax, and started to take care of Poland's interests. The state of the economy speaks very well for the effectiveness of our activity. Against us are those who now have lost these billions, remarked Jacek Kurzepa, a Law and Justice (PiS) member of Parliament (radiomaryja.pl).

It isn't possible to prepare, in a few minutes or even in several hours, tens of thousands of banners and billboards, which suddenly appear in places that need to be purchased in advance, such as in underground passages. Although I don't want to be considered someone who does not respect citizens who are fighting for their rights, some of these people probably do not know what they are protesting against. We see here the whole information war which has been going on in our country for several months,' explained Dr. Łukasz Kister (telewizjarepublika.pl).

"Astroturfing: the action against manipulation in the case of the courts has itself been manipulation" (onet.pl).

It is interesting that in the opinion of experts who have concerned themselves with marketing and disinformation activity in the Internet, the campaign warning against astroturfing in Poland has been, de facto, such an activity itself (tvn.24.pl).

When thousands of people go out on the streets to protest against the reform of the courts proclaimed by PiS, the right-wing media and public television of Jacek Kurski (TVP head) promote a new hypothesis. They try to prove that Poles were not at all going out on the streets spontaneously. According to commentators favourable to the government, the protesters have been subjected to 'astroturfing'.

What is going on? And what error are supporters of this theory making? (fakt.pl).

It is in fact a given interest group which needs to create this grassroots initiative and ensure that it is presented in such a way as to come off as genuinely arising from society. And in the opinion of the right-wing media, this is exactly what the street protests are about, not at all the idea of ordinary people, but a planned action intended to discredit the current government in the eyes of Poles. To support his suspicion that the protest was an act of political marketing and astroturfing, the commentator Robert Tekieli has taken the fact that the protesters held identical candles.

The question arises today how this protest against the court reform was prepared, in light of the fact that everyone is holding the same candles in their hands, Tekieli wondered (se.pl.).

It is worth investigating the narrative about how these events were characterized in media representing both political options. It should be noted that it is difficult to talk about an objective message, as in the pro-government media the position of the other side was unlikely to appear, except in the context of a negative comment.

In media with ties to the governing camp, journalists pointed to the purpose of the action, which was to scuttle the court reform in Poland. They pointed to the protests as almost certainly controlled. Their narrative repeated words such as “manipulation”, “disinformation”, “control”, “lobbying”, “information war” and “hybrid war”. They attributed the responsibility for this astroturfing to three organizations (related to the Opposition): “Akcja Demokracja” (Democratic Action), “Obywatele RP” (Citizens of the Republic of Poland) and the “Otwarty dialog” (Open dialogue) foundation. They noted the ties that its chairman, Bartosz Kramka, had with the businessman and philanthropist George Soros, who was said to be financing the protests. They went as far as to point to the Rossmann chain of shops as a supporter of the demonstrations. As evidence that they were organized and not a grass-roots movement, they cited the fact that the participants in the “chain of light” all held the same candles. They argued, further, that the banners, signs and high-quality technical support pointed to a controlled action. These arguments were quoted by an investigative journalist for TVP, Anita Gargas (tvp.pl). Rossmann officially denied the report that it had financed the protests.

The reaction on the other side of the political stage was marked by an irreverent tone, mostly scoffing what they called the right-wing media’s needless suspicion and searching for conspiracy theories. Journalists not associated with the government side accused the media and pro-government Internet users of web astroturfing. Beata Biel’s text, “Bot people and Agent Zero: All Poland is sniffing out astroturfing”, appeared on TVN 24’s web page on August 29 (Biel 2017).

All the texts about the described events are clearly of a tabloid nature (Piontek, 62). Tabloid-style communication is characterized by emotionalism in news, the manipulation of reality to highlight contrasts, and a tendency to be dramatic. Tabloidization, and the concept of “infotainment” associated with it, have in some way taken over today’s media. This is also a manifestation of the *diktat* of the receiver: when a mass

audience can choose content, it usually opts for entertainment, which explains its irreverent tone. If I am looking for information, it is rather sketchy and easy to understand.

Reports

To support her thesis, the author recalls a report by Ben Nimmo of the Digital Forensic Research Lab (DFRLab), which noted that bots were involved in the spreading of Twitter messages containing the hashtags #AstroTurfing, #StopAstroTurfing, and #StopNGOSoros (DFRLab 2017).

In the case described here, tweets and hashtags such as “Sprzeciwiam się wykorzystywaniu mechanizmu #AstroTurfing aby manipulować Polską!” (“I oppose the use of the mechanism of #AstroTurfing to manipulate Poland”) are cited, as well as the memes below (Figure 1).

Figure 1: Meme used on Twitter

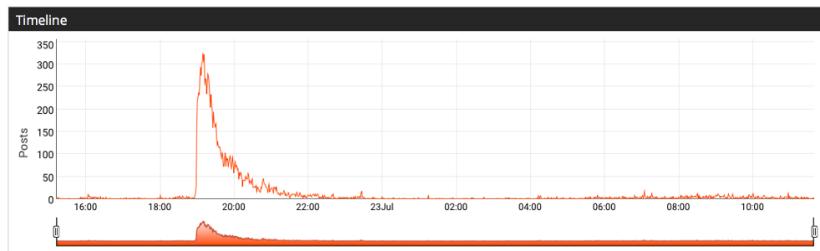


Source: https://twitter.com/nazimno_/status/888778829489156096/photo/1, accessed 10.10.2017.

(This is not a revolt, but co-ordinated and paid political marketing activity #AstroTurfing. In the USA and the UK, it is forbidden. In Poland, we bear witness to one of the most massive actions of this type in history.)

The report's author notes that after an analysis of the frequency with which this appeared, after its first appearance on July 22 it was posted on Twitter 15,882 times by 2,408 users. Nimmo notes that the Twitter activity was not constant, but rose sharply at around 9 p.m. Polish time, from a rate of three tweets a minute to over 200. The intense tweeting with this hashtag did not last even for an hour. The frequency of the posting dropped to ten per minute. Below, I have presented the tweet frequency indicated in the report, which shows a rapid increase at around 7 p.m. UTC (in Poland, 9 p.m.).

Chart 1. #StopAstroTursng and #StopNGOSoros tweets, over time, 22-23 July, 2017



Source: <https://medium.com/dfrlab/polish-astroturfers-attack-astroturfing-743cf602200>, accessed 10.10.2017.

Nimmo states that such Twitter action points to the use of a bot, and not “organic” activity, as each user who used the mentioned hashtags did so an average of 6.6 times. This is almost certainly an anomaly, given that the typical engaged user posts 3.5 times on the site. Nimmo describes some users of the service as hyperactive, as during two hours shown on the graph they posted 1,804 tweets with the above-mentioned hashtags. Nimmo concludes that this works out to one tweet in four seconds. In the researcher's view, the hyperactivity of these accounts would point to the use of bots (DFRLab 2017).

In addition to the DFRLab analysis, two further reports appeared that analyzed this case. In the first, the company DBMS presented data obtained from Google Trends that showed that between mid-June and

July 23, 2017, certain words appeared with increasing intensity in Internet searches: "protests", "constitution" and "democracy". This interest peaked on July 22 and 23. The unnamed author of the report also checked interest in the words "3Xweto" (three times veto), "łańcuch światła" (chain of light) and "PolandDefendDemocracy", finding a similar result.

The author also analyzed the level of sentiment and the number of mentions in social media using a tool called sentione, which takes a comprehensive approach to the analysis of web content. The data collection and analysis methods used for this purpose automatically reveal invisible patterns and connections in large content sets. (DBMS 2017) The analysis showed an increase in interest of over 50 per cent in the subjects of protests, democracy and the constitution. The author emphasises the higher intensity of Facebook discussions, and a high percentage of negative comments. While all the above observations can be explained by the popularity of the subjects and their connection to the strongly polarizing political scene in Poland, the following observations of the author seem interesting:

- 27% of the posters' active accounts were from outside Poland's borders;
- 75% of the posters' accounts were active (hosting at least one a month);
- 21% of the accounts taking part in the discussion had no connection to profiles in Poland;
- Poster profiles were in no way connected to networks of friends in Poland (DBMS 2017).

The fact that the studied profiles had no connections in Poland, but came from outside the country, does not yet indicate the use of bots, but such a strong possibility has arisen (Ferrara et al., 99).

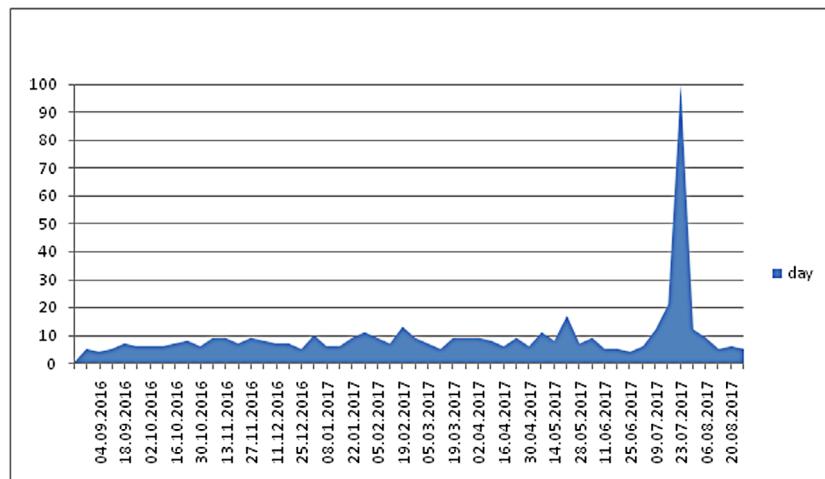
As mentioned above, yet another report was prepared by the consulting agency Apostolowie Opinii. In this case, the report's authors did not specify their data analysis tools, except for Google Trends. They pointed out that a very large accumulation of comments appeared at the same time on TVN 24's Facebook page, and a large number of new posts, observing politicians from the PO opposition party, appeared on Twitter (Apostolowie Opinii 2017). The report's authors analysed the

hashtag #astroturfing. Some of their findings concur with Ben Nimmo's observations about the times during which the greatest number of mentions occurred in social media.

This research should be of some help in assessing the analyzed events. As Bartłomiej Łódzki states, because it is possible to research quickly data obtained from the Internet, repeatedly by various institutions, think tanks and companies, representative results can be obtained. The techniques and tools used in these processes may give a very representative result (Łódzki 2017, 142).

The chart below presents the author's research, conducted with the help of the Google Trends tool, allowing quick access to the content in which Internet users took an interest, as well as their activity according to time, territory and region.

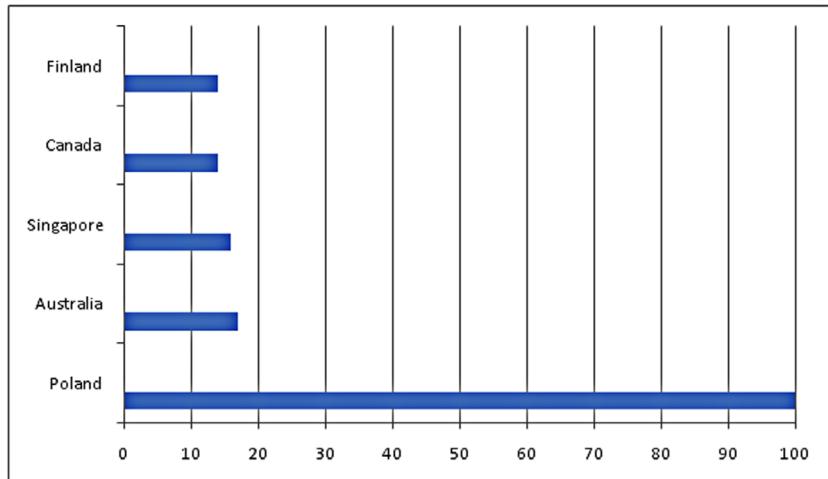
Chart 2. Interest in the subject of “astroturfing”



Source: a study on the basis of data obtained from Google Trends.

The above data present how often the term was searched with reference to its highest point in a given time and region. The value of 100 signifies the highest level of popularity of the term. A value of 50 signifies that the term's popularity was less by half. On the other hand, a zero value reports that the term's popularity amounted to less than one per cent of its highest level.

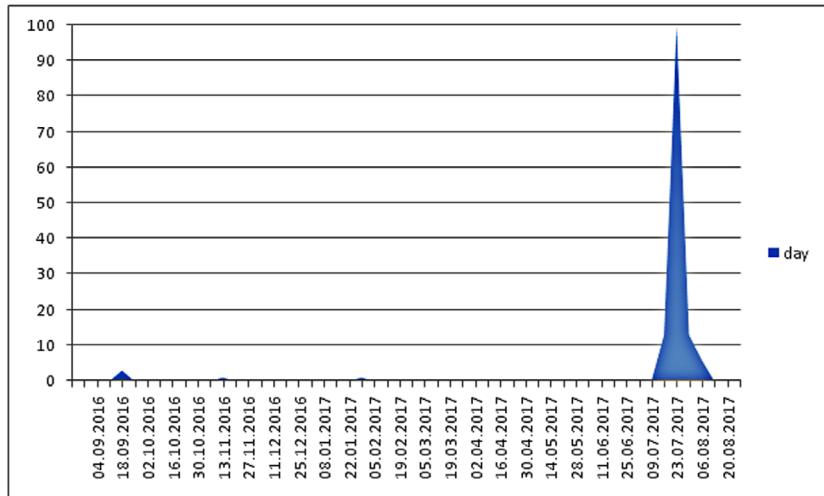
Chart 3. Interest in the term, according to country



Source: data obtained from Google Trends.

Values on the above chart are plotted on a scale from 0 to 100, where 100 represents the location in which the term was most searched, and 50 means the lower popularity of the search at this location. On the other hand, a value of 0 characterizes the place where the popularity of the term amounted to less than one per cent of its highest value. In Chart number 3, interest in the term is shown at a given period of time, varying by country. In this case the highest interest in the term was experienced in Poland. The chart below shows, additionally, that search queries for the term “astroturfing” reached their peak on July 22 and 23.

Chart 4. Interest in the term “astroturfing” in Poland



Source: data obtained from Google Trends.

Similar data were found on the growth of interest on July 22 and 23 2017, through the application of the same tool, in the terms “3xweto” (three times veto), “łańcuch światła” (chain of light) “PolandDefendDemocracy”, “konstytucja” (constitution), “wolne sądy” (free courts) and “demokracja” (democracy). This is easily explained, as they were the most prominent subjects of discussion at that time in the mass media. At that, the term “astroturfing” is specialized, referring to a specific phenomenon in the sphere of marketing; further, it is English, and thus it is debateable whether there should have been such high interest in the term on the territory of Poland. The increase in Google searches coincides, in terms of time, with the likely astroturfing campaign. Ben Nimmo’s research clearly points to the presence of bots on Twitter, while similar conclusions from the analysis of Facebook allows us to draw the conclusions that DMBS reached. It is difficult, however, to identify the source of the campaign. Nimmo speculates that it was Russia, while DMBS only enigmatically suggests a foreign source, not specifying its origin. In media discussions, three sources have been pointed to: George Soros, Russia and Germany.

Conclusions

From an analysis of both the course of the campaign and its effects, it would seem that we were dealing with professional activity. Data, produced with the help of analytical tools, point to this. However, it is difficult to identify the source. As a result of this activity, the political situation has become even more polarised, and it is difficult to foresee dialogue between the governing and opposition camps. This is exemplified on social media such as Twitter and Facebook. Chaos arose in the area of information, and it was difficult for the typical receiver of information to understand who was right. This strengthened suspicions that a conspiracy was at work. Both sides in Poland's political conflict have exploited this, as they may conclude from such campaigns that they strengthen mistrust of adversaries. The suspicion that marks relations with the European Union, deepened by accusations against Germany, is also on the hands of Eurosceptic forces in Poland. Russia, which may depend on Poland's destabilisation and isolation from the rest of Europe, also appears to be a beneficiary of this campaign.

References:

- Apostołowie Opinii, <https://www.tysol.pl/a9655--Analiza-Polityka-w-sieci-%23Astroturfing-w-mediach-spolecznosciowych-21-24-07-2017>, accessed 01.09.2017.
- Astroturfing. Akcja przeciwko manipulacji w sprawie sądów sama była manipulacją*, <http://wiadomosci.onet.pl/kraj/astroturfing-akcja-przeciwkomomanipulacji-w-sprawie-sadow-samabyla-manipulacja/lt3t2q0>, accessed 30.08.2017.
- Astroturfing. Do Polski wszedł bardzo niebezpieczny przeciwnik, który nie zjawił się tu przypadkiem*, <http://niezalezna.pl/102770-astroturfing-do-polski-wszedl-bardzo-niebezpieczny-przeciwnik-ktry-nie-zjawil-sie-tu-przypadkiem>, accessed 25.08.2017.
- Astroturfing? Po jednej i drugiej stronie z pewnością walczyły boty. Ale internet nie zmieni władzy*, <http://wyborcza.pl/7,75400,22192659,astroturfing-pojednej-i-drugiej-stronie-z-pewnoscia.html>, accessed 29.08.2017.
- Biel B., *Ludzie boty i agent zero. Cała Polska na tropie astroturfingu*, <http://www.tvn24.pl/magazyn-tvn24/ludzie-boty-i-agent-zero-cala-polska-na-tropie-astroturfingu,111,2048>, accessed 25.08.2017.
- Bernatis V. (2014), *The Taliban and Twitter: Tactical Reporting and Strategic Messaging*, "Perspectives on Terrorism", 8(6).

- Bowen S.A., Stack D.W. (2014), *Understanding the Ethical and Research Implication in Social Media*, in: DiStaso M.W., Sevick D. (eds.), *Ethical Practice of Social Media*, in *Public Relations*, New York: Routledge, pp. 217-235.
- Bu Z., Xia Z., Jiandong Wang J. (2013), *A Sock Puppet Detection Algorithm on Virtual Spaces*, "Knowledge-Based Systems", 37, pp. 366-377.
- Centrum Badania Opinii Społecznej (2017), *Media Online*, 53/2017, Warszawa, CBOS.
- Celem głównym protestów jest zmiana sytuacji gospodarczo-społecznej w Polsce, <http://telewizjarepublika.pl/dr-kister-celemglownym-protestow-jest-zmianasytuacjigospodarczo-spolecznej-wpolscie,51934.html>, accessed 25.08.2017.
- Darmowa pomoc marketingowa dla spanikowanej opozycji, <http://wpolityce.pl/polityka/349879-darmowa-pomoc-marketingowa-dla-spanikowanej-opozycji-petrusi-schetyna-siegaaja-po-najlepszych-macherow-od-manipulacji-na-pokladzie-bierzynski-i-benke>, accessed 25.08.2017.
- Delli Carpini M.X., Keeter S. (1996), *What Americans Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.
- Digital Forensic Research Lab, <https://medium.com/dfrlab/polish-astroturfers-attack-astroturfing-743cf602200>, accessed 01.09.2017.
- DBMS, *Protest zapisany w danych*, <https://edbms.pl/protest-zapisany-w-danych-astroturfing-czy-spontaniczna-obrona-demokracji/>, accessed 01.09.2017.
- Ferrara E., Varol O., Davis C., Menczer F. & Flammini A. (2016), *The Rise of Social Bots*, "Communications of the ACM", 59 (7), pp. 96-104.
- Gromadzki G. (2018), *Polska w Unii Europejskiej. Nieunikniona marginalizacja z małym znakiem zapytania*, <http://library.fes.de/pdf-files/bueros/warschau/14180.pdf>, accessed 01.09.2017.
- Hille S. & Bakker P. (2014), *Engaging the Social News User*, "Journalism Practice", 8 (5), pp. 563-572.
- Howard P.N., Kollanyi B. (2016), *Bots, #StrongerIn, and #Brexit: Computational Propaganda During the UK-EU Referendum*, www.politicalbots.org, accessed 25.07.2017.
- Jenkins H. (2006), *Fans, Bloggers, and Gamers: Exploring Participatory Culture*, New York: New York University Press.
- Klotz R.J. (2007), *Internet Campaigning for Grassroots and Astroturf Support*, "Social science Computer Review", 25 (1), pp. 3-12.
- Kovic M., Rauchfleisch A., Sele M. (2016), *Digital Astroturfing: Definition, Typology, and Countermeasures*, <https://zipar.org/publications/>, accessed 25.07.2017.
- Łódzki B. (2017), *Narzędzia internetowe w transnarodowych badaniach mediodziewczych*, in: K. Kopecka-Piech, K. Wasilewski (eds.), *Lokalne - Regionalne - Transnarodowe. Rola mediów w kształtowaniu wspólnot*, Gorzów Wielkopolski: Polska w Świecie, pp. 139-157.

- Mackie G. (2009), *Astroturfing Infotopia*, "Theoria: A Journal of Social and Political Theory", 56 (119), pp. 30-56.
- McNutt J., Boland K. (2007), *Astroturf, Technology and the Future of Community Mobilization: Implications for Nonprofit Theory*, "The Journal of Sociology & Social Welfare", 34 (3), pp. 165-178.
- McNutt J.G. (2010), *Researching Advocacy Groups: Internet Sources for Research about Public Interest Groups and Social Movement Organizations*, "Journal of Policy Practice", 9(3-4), pp. 308-31.
- Newman N. (2017), *Journalism, Media and Technology Trends and Predictions 2017*, Oxford: Reuters Institute.
- Nowoczesna i PO w manifestacjach przeciw ustawom o KRS i Sądzie Najwyższym, <http://www.wirtualnemedia.pl/artykul/nowoczesna-i-po-w-manifestacjach-przeciw-ustawom-o-krs-i-sadzie-najwyzszym-nie-pomagaja-nam-marketingowcy-benke-bierzynski-i-ciszewski-teoria-astroturfingu-to-bzdury>, accessed 25.08.2017.
- Onet i wp.pl informacyjnymi liderami portali, <http://www.wirtualnemedia.pl/artykul/onet-i-wp-pl-informacyjnymi-liderami-portali-interia-pl-blizej-gazeta-pl>, accessed 25.07.2017.
- Peeples Massengill R. (2013), *Wal-Mart Wars: Moral Populism in the Twenty-First Century*, New York: NY University Press.
- Piontek D. (2011), *Komunikowanie polityczne i kultura popularna, Tabloidyzacja informacji o polityce*, Poznań: UAM.
- Polska jest dzisiaj obiektem zmasowanej akcji dezinformacyjnej, <http://www.tvp.info/33327183/polska-jest-dzisiaj-obiektem-zmasowanej-akcji-dezinformacyjnej>, accessed 25.08.2017.
- Prior M. (2007), *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*, New York: Cambridge University Press.
- Serwisy informacyjne zyskały, <http://www.wirtualnemedia.pl/artykul/servisy-informacyjne-zyskaly-w-dniach-manifestacji-wp-przed-onetem-i-tvn24-pl-gazeta-pl-przed-interia-pl>, accessed 25.08.2017.
- Sweetser K.D. (2014), *Ethics, Reputation and Social Media*. in: DiStaso M.W, Sevick D. *Ethical Practice of Social Media in Public Relations*, New York: Routledge, pp. 205-217.
- Szeroka kampania dezinformacyjna?, <http://www.radiomaryja.pl/informacje/u-nas-j-kurzepa-organizacja-manifestacji-wsreformy-sadownictwa-stoi-poteznelobby/>, accessed 25.08.2017.
- Teoria spiskowa prawicy, <http://www.fakt.pl/wydarzenia/polityka/teoria-spiskowa-prawicywymyslili-astroturfing/77lp4w8>, accessed 30.08.2017.

UJAWNIAKY. 16 kroków, które mają sparaliżować Polskę i odsunąć PiS. Inżynierowie zborowej histerii mają chytry plan: „Wylaczmy rzad!”, <https://wpolityce.pl/polityka/349913-ujawniamy-16-krokow-ktore-maja-sparalizowac-polske-i-odsunac-pis-inzynierowie-zbiorowej-histerii-maja-tajny-plan-wylaczmy-rzad?iframe=true&width=100%&height=100%>, accessed 25.08.2017.

Vergani M. (2014), *Rethinking Grassroots Campaigners in the Digital Media: The ‘Grassroots Orchestra’ in Italy*, “Australian Journal of Political Science”, 49 (2), pp. 237-251.

Zaller J. (2003), *A New Standard of News Quality: Burglar Alarms for the Monitorial Citizen*, “Political Communication”, 20 (2), pp. 109-130.

Zhang J., Carpenter D., Ko M. (2013), *Online Astroturfing: A Theoretical Perspective*. AMCIS 2013 Proceedings.

Zheng X., Zeng Z., Chen Z., Yu Y., Rong C. (2015), *Detecting Spammers on Social Networks*, “Neurocomputing”, 159, pp. 27-34.

Where Does Creative Class Meet Trendsetters?

„The history of human creativity and of human progress is intimately intertwined with that of cities.”

(Florida, 3218-3219)

Abstract: This study aims to discuss the idea of creative cities being potential places where trends are emerging and growing, based on the arguments presented by Henrik Vejlgård in the book “Anatomy of Trend” and Richard Florida in “The Rise of the Creative Class, revisited”. Selected factors related to the place, conducive to trendsetting and the factors making cities clusters for the creative class have been compared. Will the trendsetters and representatives of the creative class meet in the same city, live in the neighborhood, go to the same event and drink coffee in the same cafe? Will these features of the cities that will attract the creative class be a magnet for trendsetters?

Keywords: trend, trendsetter, Creative Class, city

Introduction

The study of trends allows us to get a picture of changes that affect consumer choices, preferences and behaviors as a result of constant changes taking place in the market environment (Zalega 2013, 4). Understanding their “anatomy” and, above all, the process of spreading on a large scale allows for making effective business decisions. An important ele-

ment in predicting trends is the awareness of the place where they are born and from where they spread to the whole country or the world. We are also currently observing the growth of the social structure, called by Richard Florida a “Creative Class”. The distinguishing characteristic of the Creative Class is that its members engage in work whose function is to “create meaningful new forms” (Florida 2012, 911-912) and tend to cluster in urban communities, where their creativity can take root and flourish (Florida 2012, 3039).

The relationship between creativity and the emergence of consumer trends is unquestionable. But are the cities, where the current Creative Class prefers to re-locate and live, really a natural breeding ground for trends, and what is very important, spreading on the scale of the country and globally?

How can the characteristics of cities currently described as creative contribute to the emergence and spread of consumer trends? Do creative cities really have such potential? Are their resources and specificity effective and sufficient?

In the first part, we will face the phenomenon of the trendsetters, “place of action” and the Creative Class. In the second part, we will try to define the conditions that make the city both a friendly “seedbed” of trendsetters and attractive for Creative Class.

Trendsetters, “place of action” and the Creative Class

Vejlgaard talks about the trends in a context of change in style, design, what we eat and drink, what we watch, we read – in one word, in all areas where we use the word “taste” to describe our preferences (Vejlgaard 2008, 27). The life cycle of the trend involves at the subsequent stages of its development various “players”, but the key to its creation and dissemination are two groups: *trend creators* and *trendsetters*. The easiest way is to say this: trend creators create something new, while trendsetters are the first to accept the novelty and promote it (Vejlgaard 2008, 55). Who are the creators of trends and trendsetters? According to Henrik Vejlgaard, trend creators are a heterogeneous and extremely small group, composed of people of very different professions, especially creative and innovative (Vejlgaard 2008, 81). Trendsetters, on the other hand, are a much larger

group. Young people (under 30), designers, artists, rich people, homosexual men, celebrities, representatives of subcultures with high style consciousness (e.g. hipsters, body-builders, chefs, drag queens). Particularly characteristic for trendsetters is courage and openness to new products, combined with a developed sense of aesthetics. The lifestyle is also decisive, with the trendsetter personality rather than its financial resources. Trendsetters can come from different backgrounds and perform various professions. They can be artists, but also sellers, they can be homo- or heterosexual, rich or poor, old or young (Velgaard 2008, 94). They are not better educated or wealthier than most people (Velgaard 2008, 217).

Contemporary social changes have caused a change in the processes of spreading trends. Until World War II, trends flowed from rich people down the social ladder, now they are also spreading in a process that can be called "boiling". A series of examples proves that changes in styles and tastes are born among poor people, outcasts and subcultures (Vejlgaard 2008, 211).

Being a trendsetter means that a given style or taste is taken before others do it. In fact, trendsetters are buying products in a new style more often than others, but this process can also take place bypassing financial transactions (e.g., hair growth or colouring, sun avoidance due to fashion for unburied body, etc.) (Velgaard 2008, 200).

The exchange and spreading of trends obviously requires polisocial connections. Usually, trendsetters prefer to be surrounded by other trendsetters, but because they come from very different social groups, there is an exchange between different groups of people. They do not need to know each other or talk; it's enough that they can watch each other. Where do such contacts come from? Most often in the places conducive to observation and drawing inspiration: at fashion shows, concerts, theatre performances, film premières, at festivals, art exhibitions and auctions, industry events. They have also connection to mainstream as most of their social group members (young people, gays or artists) belong to mainstream.

Observing these trendsetter groups and events in which they participate, we can notice how often there are contacts between seemingly distant worlds – writers and musicians participate in fashion shows, celebrities spend time with rich entrepreneurs, homosexuals are interested in design, designers draw inspiration from street fashion etc. (Velgaard 2008, 81-83).

Where are the best chances to meet many different people in a short period of time, to participate in cultural and artistic events, to watch emerging trends, sometimes even without moving from your own street (Vejlgaard 2008, 217)?

More or less since the 13th century, there have been successive economic structures that gather society around them. The centre of this structure is the dominant city, accumulating significant financial, technical, cultural and ideological power, though not necessarily political (Attali 1991, 24). Jacques Attali points to cities such as: 1300 Bruges, 1450 Venice, 1500 Antwerp, 1550 Genoa, 1650 Amsterdam, 1750 London, 1850 Boston, 1930 New York (Attali 1991, 27).

Trends arise and grow in metropolitan centres in which there are polisocial groups (Vejlgaard 2008, 117).

There is a limited number of cities with “trendsetting potential” (Vejlgaard 2008, 159). Here are some crucial conditions that must be met to get this potential:

- high diversity level among the trendsetters,
- very good connection with the world (like flight connections),
- repeated type of trend born in the city in the past (like we can see it in Paris, London, Milan, Tokyo, Los Angeles, San Francisco, New York) (Vejlgaard 2008, 160).

In the process of spreading trends, the media currently play a great role (Vejlgaard, 2008 215), but even the current transfer of many activities and communication to the Internet does not change the fact that only the combination of virtual reality with real contact gives the intended trendsetting effect. *Together the smartphone and the media platform such as Instagram constitute an architecture that enables spaces as clubs, cultural events and other locations where bodies and web are connected, smartphones appear to become sites where affect is released, channelled and directed.* (Brodmerkel, Carah 2016, 140).

Business Insider has published a report on places in USA best for trendsetters, taking into account such factors as: access to public transport, a large number of bars and restaurants, happening night life and great entertainment. “And the winner is”:

1. Manhattan, NY
2. San Francisco, CA
3. Chicago, IL
4. Los Angeles, CA
5. Seattle, WA
6. Brookline, MA
7. Arlington, VA
8. San Diego, CA
9. Portland, OR
10. Mountain View, CA (Pollard 2012).

It is estimated that in 2016, 54.5% of the world's population lived in urban settlements and it is predicted that by 2030, urban areas will be inhabited by 60 percent of the world's population (UN Data booklet, 2016, ii).

From the point of view of the economic growth, the city has always been of great importance – as we have already said, since the Renaissance era we have been dealing with cultural and economic centres, characterized above all by favourable communication (Attali 1991, 27). Currently, their importance is growing, mainly due to the current knowledge-driven economy. Cities, however, are not just the storage of smart people, but rather a structure that allows connections, network building and consummation of this innovative combinations (Florida 2012, 3074). Professor Edward L. Glaeser from Harvard University said: *Globalization and new technologies attract people to big cities, by increasing the returns to urban proximity. (...) Globalization and technological change have increased the returns to being smart; human beings are a social species that get smart by hanging around smart people* (NY Times).

History shows that for centuries creative creators gathered in places that provided them with contact with other creative people, created an appropriate environment and support for creative activities. This kind of support is provided by cities (Florida 2012, 3221).

Florida dealt with a phenomenon he called the geography of the *Creative Class* (Florida 2012, 3291). Who is the Creative Class?

He defines the Creative Class by the occupations that people have and he distinguishes two main components: Super-Creative Core of the Creative Class (scientists and engineers, university professors, poets and novelists, artists, entertainers, actors, designers, and architects, as well as the thought leadership of modern society: non-fiction writers, editors, cultural figures, think-tank researchers, analysts, and other opinion makers. *I define the highest order of creative work as producing new forms or designs that are readily transferable and widely useful-such as designing a consumer product that can be manufactured and sold; coming up with a theorem or strategy that can be applied in many cases; or composing music that can be performed again and again.* This kind of work they perform regularly and it's what they are paid for. Second group includes "creative professionals" who work in a wide range of knowledge-intensive industries, such as high-tech, financial services, the legal and health care professions, and business management. *These people engage in creative problem solving, drawing on complex bodies of knowledge to solve specific problems.* Their jobs usually require a high degree of formal education. They can also move up to the Super-Creative Core: *producing transferable, widely usable new forms is now their primary function* (Florida 2012, 911-927). From the very definition of the creative class, we can see that there are potential creators of trends among them as well as their popularizers.

Creative urbanity

As we can see, the number of cities and their importance is increasing globally, and both trendsetters and the creative class prefer urban environment. Let's look whether what encourages development of the Creative Class is also attractive for trendsetters.

Creative people do not move for traditional reasons (Florida 2012, 3031).

They look for rich, high-quality facilities and experiences, openness to the diversity of all species, and above all, to confirm their value as creative people (Florida 2012, 3031). Cities stimulate creativity (Florida 2012, 3269). *Cities have long functioned as critical containers and mobilizers of creativity, attracting creative people from the surround-*

ing countryside while providing the structures (Florida 2012, 3224-3225). As Jane Jacobs, the urban theorist said, cities generate diversity of every kind because of so many people being so close together and bringing *so many different tastes, skills, needs, supplies* (Florida 2012, 3104-3108).

What can people be paying Manhattan and Chicago rents for, if not for being near other people? (Lucas 1988, 39).

As the key to understanding geography of the creativity, Florida defines three factors of economic development, the so-called 3 T's: technology, talent and tolerance (3594). 3 T's comprise the Creativity Index (3598).

In the context of Maslow's needs pyramid, economic growth, influenced by all three elements of 3 T's together, contributes to the growth of the society's wealth, and thus greater interest in changes in styles and tastes, and hence greater efficiency of trends spreading by trend-setters.

Because the trends we are talking about can not ignore technological development, coexistence with the other T – technology and talent, apart from the economic factor, also gives the opportunity to be inspired by the development of knowledge and the way of thinking of "smart people".

And the tolerance as the crucial factor for the creation of a polisocial environment, which is one of the keys for successful trendsetting. Places characterized by tolerance, i.e. openness to diversity, attract talented people, thus expanding their technological capacity and creativity, and thus achieve a significant economic advantage (Florida 2012, 3649).

As we already know, trendsetters are characterized by great diversity, and their effectiveness requires a multicultural environment in which they function. That is why openness to diversity in a given city determines whether trendsetters will appear in it or not.

What does tolerance look like in practice? First of all, there are no entry barriers and fast acceptance of newcomers (Florida 2012, 3663). It is ethnic and national diversity, multiculturalism. For trendsetters, it is primarily an inspiration flowing from the richness of other cultures, especially interesting in the context of many different cultures, not only coexisting but also inter-penetrating each other.

Florida perceives a vibrant gay community as a signal of openness to different types of people, and as a factor attracting not only immigrants and national minorities, but also *eggheads, eccentrics, and all the other non-white-bread types who are the sources of new ideas* (Florida

2012, 3722-3723). *Where gay households abound, geeks follow* (Bishop, Florida, 3723-3724).

Representatives of the creative class prefer authentic, indigenous or organic meeting places with a wide offer, which they will also be able to influence (Florida 2012, 2589-2591). What is important for Creative Class is the so-called street-level culture. Easy to reach, eclectic, unlimited hours of the day or business hours. Giving you the freedom to observe and interact with people in exotic ways: hippies, young people with shocking hairstyles, foreigners in specific outfits, etc. Where you can be with musicians, painters, writers even without being an artist. Local, because created by people living in the area. Street, with a tendency to focus along the street with lots of small shops, bookstores, cafes, art galleries, and various hybrid spaces. Living day and night, giving creative representatives the opportunity to contact culture on a daily basis, without having to match an intense schedule to the hours of work of large cultural institutions (Florida 2012, 2477-2574).

The eclectic scene is also an important factor attracting and connecting different communities (Florida 2012, 2477-2574), what is crucial for polisocial connections, so crucial for trendsetters efficiency (Vejlgaard 2008, 117).

Big Asian cities are the largest and densest in the world, but their innovative and creative impetus pales in comparison to places like New York and London in finance, Paris in fashion, Milan in design or Silicon Valley in technology. The advantage for cities lies in their pedestrian-friendly scale, which includes buildings of various purpose and average altitude, numerous bars, cafes and other third places, and active street life that facilitates interaction between people (Florida 2012, 5022-5025).

Mets provide also access to higher-level events and cultural institutions: they have both a base (theatres, museums, concert venues) and resources for organizing expensive concerts or events. They are also stimulating creativity, even if not on everyday level. And what important: they are places and occasions, where trendsetters from different social groups can meet, watch and get inspiration (Vejlgaard 2008, 83). But the street level culture is a must (Florida 2012, 2523-2524). For both: Creative Class and trendsetters.

The bohemians and artists milieu, being inspiration for the creative class and breeding ground for trendsetters, can, however, make the price increase of the property to which it contributes, an obstacle to its

own inspirational activity. *In New York City, the old refrain goes a little something like this: Artists move into a neighbourhood, make it cool, get pushed out by rising housing costs as area desirability increases, artists find another place to settle – and on that cycle goes* (Rosenberg no date). Surprisingly, it doesn't mean the artists leave the city: *As of 2017, at a time when rents dance around their apex, there are more artists living in New York City than ever before* (Rosenberg no date).

Although in the case of the trendsetters, neither direct contacts nor networking are crucial for their effectiveness, the social structure of creative cities can inspire and supports them. In the process of spreading trends, people do not have to communicate or establish relationships – the whole process is based rather on circulation, mixing and observation (Vejlgaard 2008, 215). It is not about stable or temporary contacts – contacts may not be needed at all.

There are other connections between Creative Class and trendsetters: Renfrow in his work “Open City” examined the concentration of people with the personality of “open-to-experience” in metropolitan cities and obtained a result that confirmed that a large number of people with the personality of “open-to-experience” appears in cities that have also a significant Creative Class indicator: San Francisco, Los Angeles, Austin, New York City and San Diego (Renfrow, Florida 2012, 3870-3874). Interesting, that most of this cities are also mentioned by Velgaard as “place of action” for trend creation and trendsetting.

Twentieth century sociologist Robert Park made a very interesting observation: *Great cities, have always been melting pots of races and of cultures. Out of the vivid and subtle interactions of which they have been the centres, there have come the newer breeds and the newer social types. They have multiplied the opportunities for the individual man for contact and association with his fellows, but they have made these contacts and associations more transitory and less stable* (Florida, 3247-3248). Cities attract with their anonymity, loose relationships, built mainly in the circle of friends than with family or neighbours.

Because trends and trendsetting are about novelty, this transitory of relationships in big cities, combined with their diversity and a large number of their inhabitants, creates potential for constant inspirations and observations. Possibility of loose circulation between groups, without need or pressure to establish permanent relationships, gives the opportu-

nity to contact many different groups and observation without obligation. Loose and temporary relations also mean lack of control and freedom of action and to be different. Park wrote also: *The small community often tolerates eccentricity, the city, on the contrary rewards it*, thus opposing small communities to big cities (Florida, 3261).

Conclusions

Comparing cities with the highest percentage of creative class in the USA (according to Florida) and cities considered by Vejlgaard as cities with trendsetting potential, we can see Los Angeles, San Francisco or New York on both lists. So in fact representatives of the creative class and trendsetters could meet and inspire on the same scene in mets.

Everything is about being creative and open, and that's what characterizes both Creative Class and trendsetters. It does not matter that some are rewarded for their creativity financially, while others are not necessarily; important that both need a stimulating environment with similar parameters: a large, open and diverse community, with a high degree of tolerance, enabling mutual observation, inspiration, and the penetration of groups and culture; encompassing both artistic environments as well as subcultures and young people; with access to both high art and street level culture (Florida 2012, 2504). And that's why they can meet – in creative metropolises that will take care of everything to ensure them.

References:

- Attali J. (1991), *Millenium: Winners and Losers in the Coming World Order*, New York: Times Books.
- Brodmerek S., Sarah N. (2016), *Brand Machines, Sensory Media and Calculative Culture*, London: Palgrave Macmillan.
- Florida R. (2012), *The Rise of the Creative Class-Revisited: Revised and Expanded*. Basic Books. Kindle Edition.
- Glaeser E.L., *Why Has Globalization Led to Bigger Cities?*, <https://economix.blogs.nytimes.com/2009/05/19/why-has-globalization-led-to-bigger-cities>, accessed 28.02.2017.

- Lucas R. Jr (1988), *On the Mechanics of Economic Development*, “Journal of Monetary Economics” 22, pp. 3-42.
- Polland J. (2012), *The 10 Best Cities For The Young And Trendy*, <http://www.businessinsider.com/10-best-cities-for-young-trendsetters-2012-7?IR=T>, accessed 10.02.2017.
- Rosenberg Z., *Where Do New York City's Artists Live?*, <https://ny.curbed.com/2017/7/13/15964650/where-do-new-york-citys-artists-live>, accessed 10.02.2017.
- United Nations, Department of Economic and Social Affairs, Population Division (2016), *The World's Cities in 2016 – Data Booklet* (ST/ESA/ SER.A/392).
- Veljgaard H. (2008), *Anatomia trendu*, Kraków: Oficyna Wolters Kluwer.
- Zalega T. (2016), *Nowe trendy konsumenckie jako przejaw innowacyjnych zachowań współczesnych konsumentów*, “Nierówności Społeczne a Wzrost Gospodarczy”, 46 (2).

This book deals with an issue that has global importance not only in a wide range of contemporary social sciences, but also in the daily lives of billions of people: increasingly large data sets. This plethora of data presents, at every level of processing, a serious challenge for those who need to make sense of it. For them, the skill of using this data in an appropriate, well-considered way has become essential, and demands specialist knowledge. However, there is a lack of books that can provide guidance in this new area of knowledge. The book under review fills this gap.

Dr. hab. Marcin Brocki (Jagiellonian University in Kraków, Poland)

Rapid changes in the world economy demand that those entering the labour market be flexible, possessing both high interpersonal competence and the ability to develop digital skills. (...) The book titled *Big Data, digital marketing and trendwatching*, edited by Agnieszka Dytman-Stasieńko and Agnieszka Węglińska, is intended to help in the acquisition of key abilities in the areas of digital skills, the foreseeing of trends and marketing in the digital world. (...) Both the subject and the aim of the work deserve recognition, especially in light of their novelty and of their practical implications (particularly in view of the author-practitioners who have taken part in the volume's preparation).

Dr. hab. Andżelika Kuźniar (SGH Warsaw School of Economics)



ISBN 978-83-7977-341-1
9 788379 773411>

