The Influence of Online Advertising on the Information Quality of the World-Wide Web

Andrej Cvetkovski
Mother Teresa University, Skopje, North Macedonia
Email: acvetk@gmail.com

Abstract—**The technological underpinnings that have enabled the Internet are constantly developing and improving: connectivity is ever cheaper and more accessible, increasingly efficient content-management tools are available to the general public for free, ultra-mobile handsets (smartphones) are already equipped with high-performance processors that enable web-browsing capabilities comparable to those of bigger computers. As a result, everyone can freely generate and consume digital content daily. The only catch is that the content nearly always comes served with web advertisements paid by advertisers. The ad revenues go to the content creators and maintainers in order to support the entire process. Given the size of today's Internet and the amount of time spent consuming content, paid advertisement becomes a non-negligible source of income for some content publishers and website owners. This, in turn, gives strong rise to websites constructed for the sole purpose of making profit from web advertising operations, where accuracy and quality of content are of secondary or no importance. Such websites are created and managed by people who do not have relevant qualifications or genuine interest in the field. As a result, the served content ranges from worthless, automatically generated, contrived, scrapped, plagiarized, to purposeful spreading of rumors, fake news and banal, provocative or vulgar content. This paper analyzes the mechanisms that cause the predominant data on today's web to be worthless. Also considered are mechanisms by which such unwanted phenomena could be reduced and kept under control.**

Index Terms— **online marketing, fake news, misinformation, disinformation, yellow journalism**

I. INTRODUCTION

With the advent of web hardware and software technologies as means of transmitting multimedia information and providing a very high degree of end-user interactivity and mobility, the last few years have seen new ways of social communication that are more efficient and qualitatively different from traditional ones. With its increasing availability and ubiquity, the World Wide Web has imposed new standards and business models in social networking, and thus in online marketing [1].

The proliferation of web technologies for massive use has led to an increase in the availability and accessibility of free educational information to the general public. This trend has seriously challenged the position of authors and publishers of traditional printed books and magazines as well as traditional electronic media (radio, television) and has led to the development of new models of presence and sustainability of the publishing and media businesses. This in turn led to a model of open access to cover author and distribution costs, which establishes that copyrighted information is available publicly and free of charge, but the end user must, along with the information, consume ads paid by third parties - advertisers, without a clear possibility to skip or select them [2]. Thus, in order for many authors and publishers to continue their work, they have to adopt the Internet advertising as a chief source of income and a mandatory and inevitable part of their business model.

In the context of Internet advertising, we can distinguish three main types of entities: (1) content creators such as authors, news agencies, and publishers (even though potentially any network user can appear as a creator or publisher); (2) highly visited web media (e.g. online newspapers and magazines, news portals and aggregators, multimedia websites and social networks etc.); and (3) advertisers (usually including manufacturers and traders or other suppliers of goods and services) who want to do paid marketing by sending promotional commercials to a broad or targeted audience. Additionally there can be advertising agencies and advertising affiliates. The former specialize in development and use of technology to generate and send appropriate ads to a target audience, track the statistics and the success of the ad delivery and act as brokers between the content publishers and the advertisers. The latter usually do independent promotional work in the interest of the advertiser, e.g. by directing audience to the ad-serving sites etc. Each of these entities owns a share of the value flowing through this web marketing triangle: Creators create content that should be attractive and sought after; Internet media and social websites attract a large number of active users and have the immediate opportunity to send them content generated by the creators mixed with promotional messages; Advertisers and the involved agencies emit money to the publishing media as compensation for the advertising of products and services, and support content creators, thus encouraging them to create new, more engaging content for higher revenue.

But what is the effect of the existence of such a marketing triangle, i.e. what is its impact on the quality of

information whose creation and maintenance are motivated chiefly by earnings and marketing? To date, relatively little has been published on these specifics of new digital information exchange and marketing channels and their impact on the quality of available information. This paper provides a brief review of existing published research results related to the occurrence of extreme mass misinformation - fake news, rumors, conspiracy theories, hoaxes, misinformation and similar rapidly spreading nonsense in connection with the online marketing mechanisms which, driven by the idea of making money from paid advertising, cause the dominant data on today's web to be worthless content. Considered are also measures that can be applied in order to reduce and control such adverse tendencies.

II. Online Advertising and Social Networking

A key feature of modern web communications is that it enables any individual to publish data online that becomes instantly accessible worldwide. The immediacy and generality of such a system enable the existence of today's mass social networks in which each individual or competitor can, without technical knowledge, open their own websites, profiles or social network pages, in order to network with other existing individuals, pages or profiles, and comment, post positive or negative opinions and feedback, both on their own pages, and on the pages of others that allow such interaction. This can be done anonymously or using a number of real or fake identities.

The phenomenon of viral spread of information and misinformation is technically supported by the possibility of the so-called social sharing of posts according to personal findings and taste by social network users and peers. This type of sharing mechanism plays a key role in the rise and the huge impact of the social networks: some of these posts spread widely at high speeds and thus have an unexpectedly strong impact on important individuals as well as institutions, large organizations, markets or even whole countries.

Research shows that a significant portion of the planet's population receives information from social networks on a daily basis, especially in the less-developed countries [3]. Therefore, in addition to user-generated content, it is becoming more common to present direct paid advertisements and marketing messages through these networks, because such messages can easily reach a huge number of people.

In the typical web marketing model, the owner of a highly visited online media (website) appears as an intermediary between potential advertisers (possibly represented by a marketing agency) and visitors to his medium as potential users of the advertised product. The main funds come from advertisers, and are distributed among (1) the marketing agency (2) the media as content carrier and (3) content creators.

On social media, however, the situation is almost paradoxical: By default, all users of a social network join the network voluntarily, with the sole expectation of networking with friends or peers and receiving up-to-date information from their locales. The social network then sends marketing messages to its users - who are at the same time practically the only creators of content on the network. Social networks are known to be abusing personal data daily to serve precisely targeted marketing messages to their users. In this way, social network users effectively pay for the opportunity for virtual socialization with their time, privacy and voluntarily agree to be the target of a barrage of marketing or promotional messages, but as content creators, they do not receive any compensation.

III. Online Advertising and the Yellow Press

The terms yellow press and yellow journalism have been in use for more than a century now. Mott [4] defined yellow journalism in 1941 as having five distinctive elements: provocative titles, catchy graphics, pseudoscience claims and fake experts' opinions, attractive full-color supplements, and expressions of sympathy with victims of social or political injustice.

Yellow journalism lived to see its renaissance with the rise of the web and the disruptive web advertising, while retaining most of its old-day features. Its well-known products such as fake news, scandal mongering, slander and defamation of prominent people, sensationalism, publishing misinformation and disinformation continued to be used in the web era as an excellent basis for quick buildup of ad audience. Thus we have a situation in which fake news, which is considered one of the biggest problems of today's online journalism, is also one of the most fecund and frequently used platforms for online advertising.

In the modern context of web media, fake news is fabricated and distributed extensively by individuals or groups who in the real-world are most likely irrelevant, anonymous and completely unqualified, and whose sole purpose is that of attracting as wide a population as possible with cunningly concocted websites that serve paid ads. The goal is clear - by generating content that is primarily shocking, and often completely untrue and worthless, they try to maximize the number of visitors to their web media, and thus to maximize their revenue. The information quality of the content offered in this way is proportional to the qualifications of their creators as well as to the interests of the mass of uneducated audience. In the scientific literature in recent years fake news has been treated by several authors [5-24].

IV. Disinformation and Measures to Fight It

Unfortunately, many of the players in the marketing triangle have mainly pecuniary interests, and little or no interest in creating and distributing quality information or knowledge. Because of this, more and more data is generated which, speaking at the mildest is of poor quality: Content creators, in search of a bigger share of the revenue cake, often resort to fabricated information that has little touch with reality, the common sense or the fundamental sciences. Social networks, on the other hand, in search of higher profits, have a tacit interest in tolerating and even encouraging this hyper-production of misinformation and disinformation, even though

officially they have policies of no tolerance to fake news. Advertisers have an interest in pushing their content to the masses without any regard to the quality of the information used as the carrier of their advertising messages. In this ignoble triangle, despite the various offers of automatic solutions for detection and removal of disinformation, the defense of end users for the time being practically comes down to the efforts of raising awareness and honing their own common sense criteria for recognizing and ignoring media sources whose basic characteristic is profit maximization and focusing on sources that advocate the higher ideal of distributing true and useful knowledge of long-term value to humanity.

The fake news is believed to have had a key impact on the election of the 45th US president. The case of the appearance of highly visible fake news originating from North Macedonia, which had a clear and non-negligible influence on the US presidential election in 2016, is treated in detail in [16]. This paper also points out that social networks play a key role in spreading disinformation since they allow their spread across multiple routes simultaneously, through a number of anonymous or poorly visible users, without a clear reputation, who do not apply prior fact-checking, research, selectivity or editorial reasoning. In [17], the authors conclude that the degree of trust of the media that publishes information can be a useful metric for refuting rumors intended for political persuasion of the masses. However, it is also pointed out that the direct refutation of publications, i.e. labeling a medium as unreliable contributes to its visibility in the public, i.e. any mention of disinformation aids its diffusion.

Detection and suppression of misinformation and fake news are important because their spread is to the detriment of high social norms. Given the increasing speed of generating fake news and rumors, the computer science literature offers solutions for automatic detection and removal of fake news and announcements by anonymous paid publishers of defamatory content [7, 8, 11, 15, 22, 24, 25]. In [15], the authors propose a platform for collecting, detecting and analyzing online disinformation posted on microblogging platforms such as Twitter. According to their results, disinformation is ahead of factual information by an average of 10-20 hours, mainly spread by very active network users, while passive or new users passively consume this disinformation. Thus, social networks in the context of information dissemination can be called "massive online disinformation networks". A survey of the current state-of-art technologies for the detection of fake news is presented in [8] and the references therein. Generally, fake news detection systems use linguistic clues based on machine learning, network analysis or a hybrid approach. Linguistic approaches look for cues of deception in individual words or multi-words, the used syntax, or the logical contradictions and omissions of facts in the semantics of the text. Network approaches on the other hand, is based on metadata about the sources of the scrutinized data and their behavior. Currently hybrid approaches comprising both linguistic and network methods are considered most promising. In [11], the authors report three types of fake news: serious fabrications, large-scale hoaxes and humorous fakes (satire, parody, sarcastic news etc.) They propose that for better accuracy, the task of fake news detection should be separated by type. Ref. [22] considers the generalized problem of automatic removal of web noise, that is, data elements of web pages degrading the quality of the web, including irrelevant images and links, image advertisements, background images, etc. It is proposed that such removal can be done by analyzing the tags in the HTML page structure. In [24], systems have been proposed that achieve 99% success in the automatic detection of disinformation placed on Facebook only on the basis of profiling the users of the social network who liked the post. This result indicates that there is a clear dichotomy in the population of users who have a criterion for elementary life values and those who do not have such a criterion.

In [15] it is indicated that most false news is broadcast by a relatively small number of organized centers. It is an encouraging fact that social network administrators are in a position to take a decisive step in controlling and regulating the informational pollution, but their lack of interest must be balanced with external factors such as legislation. In line with this is the entry into force of European Union's General Data Protection Regulation (GDPR) for protection of citizens of the European Union. The extent to which this regulation will help regulate the risks of spreading misinformation and misuse of citizens' personal profiles for corporate and state purposes remains to be researched in the near future.

V. CONCLUSION

The new channels for data transmission and distribution, together with the high interactivity and the ability to easily create user-generated content they provide to end users enable new types of collaboration between the three main types of entities in Internet marketing: (1) content creators and publishers, (2) advertisers performing paid marketing and (3) masses of consumers acting individually or organized in large social networks. Each of these entities owns or creates value that drives the flow of disinformation: content creators create content that is attractive and sought after; advertisers issue cash as compensation for advertised products; social networks have personal data on the interests of many individuals and are in a position to aggressively distribute paid, precisely targeted marketing messages to them. Each of these players in the marketing triangle has a financial interest in broadcasting data that, at the mildest, is not quality information: Content creators, in search of a bigger share of the pie, often resort to fabricated information. In search of higher profits, social networks have a tacit interest in tolerating and even encouraging the creation and distribution of such disinformation. Advertisers have an interest in pushing their interests to the masses without regard to the quality of the information used as the carrier of their advertising messages. In this ignoble triangle, despite the offered

automatic solutions for detection and removal of disinformation, the defense of end users for the time being practically comes down to the application of raising awareness and the sharpening of their own common sense criteria for recognizing and ignoring media sources whose main purpose is profit and directing attention to sources that advocate the high ideal of distributing true and useful knowledge of long-term value to humanity.

CONFLICT OF INTEREST

The author declares no conflict of interest.

AUTHOR CONTRIBUTIONS

This paper represents author's own work.

REFERENCES

[1] R. Owen and P. Humphrey, "The structure of online marketing communication channels," *Journal of Management and Marketing Research*, vol. 2, no. 1, 2009.

[2] H. J. Rotfeld, "The stealth influence of covert marketing and much ado about what may be nothing," *Journal of Public Policy & Marketing*, vol. 27, no. 1, pp. 63-68, 2008.

[3] R. Miller and N. Lammas, "Social media and its implications for viral marketing," *Asia Pacific Public Relations Journal*, vol. 11, no. 1, pp. 1-9, 2010.

[4] F. L. Mott, "American journalism 1690-1940," vol. 4. Psychology Press, 2000.

[5] M. Balmas, "When fake news becomes real: Combined exposure to multiple news sources and political attitudes of ine-cacy, alienation, and cynicism," *Communication Research*, vol. 41, no. 3, pp. 430-454, 2014.

[6] F. Sobbrio, "Citizen-editors' endogenous information acquisition and news accuracy," *Journal of Public Economics*, vol. 113, pp. 43-53, 2014.

[7] E. Bakshy, S. Messing, and L. A. Adamic, "Exposure to ideologically diverse news and opinion on facebook," *Science*, vol. 348, no. 6239, pp. 1130-1132, 2015.

[8] N. J. Conroy, V. L. Rubin, and Y. M. Chen, "Automatic deception detection: Methods for finding fake news," in *Proc. the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1-4, 2015.

[9] M. Gentzkow, M. B Wong, and A. T. Zhang, *Ideological Bias and Trust in Social Networks,* 2015.

[10] M. Gentzkow and J. M. Shapiro, "Ideology and online news," in *Economic Analysis of the Digital Economy*, pp. 169-190, University of Chicago Press, 2015.

[11] V. L. Rubin, Y. M. Chen, and N. J. Conroy, "Deception detection for news: Three types of fakes," in *Proc. the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1-4, 2015.

[12] M. D. Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi, "The spreading of misinformation online," in *Proc. the National Academy of Sciences*, vol. 113, no. 3, pp. 554-559, 2016.

[13] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol. 59, no. 7, pp. 96-104, 2016.

[14] S. Flaxman, S. Goel, and J. M. Rao, "Filter bubbles, echo chambers, and online news consumption," *Public Opinion Quarterly*, vol. 80, no. S1, pp. 298-320, 2016.

[15] C. C. Shao, G. L. Ciampaglia, A. Flammini, and F. Menczer, "Hoaxy: A platform for tracking online misinformation," in *Proc. the 25th international conference companion on World Wide Web*, pp. 745-750. International World Wide Web Conferences Steering Committee, 2016.

[16] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 211-36, 2017.

[17] A. J Berinsky, "Rumors and health care reform: Experiments in political misinformation," *British Journal of Political Science*, vol. 47, no. 2, pp. 241-262, 2017.

[18] L. Boxell, M. Gentzkow, and J. M. Shapirom, "Greater internet use is not associated with faster growth in political polarization among us demographic groups," in *Proc. of the National Academy of Sciences*, pp. 201706588, 2017.

[19] L. Boxell, M. Gentzkow, and J. M. Shapiro, "Is the internet causing political polarization? Evidence from demographics," *Technical Report*, National Bureau of Economic Research, 2017.

[20] P. Castellani and M. Berton, "Fake news and corporate reputation: What strategies do companies adopt against false information in the media?" in *Proc. Toulon-Verona Conference" Excellence in Services,"* 2017.

[21] DJ Flynn, B. Nyhan, and J. Reifler, "The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics," *Political Psychology*, vol. 38, no. (S1), pp. 127-150, 2017.

[22] M. John and J. S. Jayasudha, "Tag based noise removal from web pages," *Journal of Theoretical & Applied Information Technology*, vol. 95, no. 22, 2017.

[23] Bartosz Redlicki. Spreading lies. 2017.

[24] E. Tacchini, G. Ballarin, M. L. D. Vedova, S. Moret, and L. D. Alfaro, Some like it hoax: Automated fake news detection in social networks. arXiv preprint arXiv:1704.07506, 2017.

[25] E. Mustafaraj and P. T. Metaxas, "The fake news spreading plague: was it preventable?" in *Proc. the 2017 ACM on Web Science Conference*, pp. 235-239. ACM, 2017.

Andrej Cvetkovski holds the Ph.D. degree in Computer Science from Boston University, Boston MA, USA, a Master's degree from New York University Tandon School of Engineering, Brooklyn, NY, USA, and a BS degree from Ss. Cyril and Methodius University, Skopje North Macedonia. He is currently Associate professor at Mother Teresa University, Skopje North Macedonia.