Censorship: A Reaction to Disinformation on the World Wide Web

Alexandre Eustáquio Perpétuo Braga*
Adilson Luiz Pinto**
Enrique Muriel-Torrado**
Moisés Lima Dutra**

ABSTRACT

The Internet and the popularization of smartphones in the 21st century have given instantaneity to information. By these same digital means –imbued with the power to influence politics, the market, culture, and health– disinformation blossoms. Censorship, understood as the suppression of content and/or suspension of users on social media, has been used as one way to combat disinformation on the Web. This drive to sanitize digital networks carries inherent risks. In the context of 'Infodemics,' investment in user's media education should be encouraged. As methodology, the phenomenon of disin-

* PGCIN/MINTER, Universidade Federal de Santa Catarina, Polícia Federal do Brasil, Brasil   alexbraga@hotmail.com
** PGCIN, Universidade Federal de Santa Catarina, Brasil   adilson.pinto@ufsc.br   enrique.muriel@ufsc.br   moises.dutra@ufsc.br

formation on the Web and the efforts to curb it were researched in both scientific literature and in Brazilian legislation. To combat disinformation, the users' education (in a sense of enabling them to filter, understand, and interpret the information that they gather) should be the main goal, as entrusting this task to third parties could bring undesirable side effects. Meanwhile, platforms, traditional media, and governments pose as 'Guardians of the Truth.' The social impacts of disinformation and of the efforts to suppress false content on the Web are discoursed in this paper through usage of MapReduce Text Mining. The paper concludes that disinformation takes on many connotations –from humorous appeal to manipulation–. As long as the users' informational competence has not been developed, both the platforms as well as governments must act to minimize the undesirable effects of this phenomenon.

Keywords: Censorship; Disinformation; MapReduce – Text Mining.

Censura: reacciones ante la desinformación en la *World Wide Web*

Alexandre Eustáquio Perpétuo Braga, Adilson Luiz Pinto, Enrique Muriel-Torrado y Moisés Lima Dutra

Resumen

Internet y la popularización de los teléfonos inteligentes en el siglo XXI han dado instantaneidad a la información. La desinformación florece a través de estos medios digitales, pues entre ellos circula poder sobre la política, el mercado, la cultura y la salud. La censura, entendida como la supresión de contenidos y/o suspensión de usuarios en las redes sociales, se ha utilizado como una manera para combatir la desinformación en la *Web*. Este impulso por 'desinfectar' las redes digitales conlleva riesgos inherentes. En el contexto de la 'infodemia' debe fomentarse la inversión en la educación de los usuarios en torno a los medios digitales. Como metodología, el fenómeno de la desinformación en la *Web* y los esfuerzos para frenarlo fueron investigados tanto en la literatura científica como en la legislación brasileña. Para combatir la desinformación, el objetivo debe ser educar a los usuarios, ya que confiar esta tarea a terceros podría traer indeseables efectos secundarios. Esto puede verse en el hecho

de que, debido a la 'infodemia', las plataformas, los medios tradicionales y los gobiernos han aprovechado para autonombrarse como 'guardianes de la verdad'. Este artículo aborda los impactos sociales de la desinformación y los esfuerzos por suprimir el contenido falso en la *Web* utilizando las técnicas de minería de texto por MapReduce. El estudio concluye que la desinformación adquiere muchas connotaciones, desde el atractivo humorístico hasta la manipulación. En tanto que la competencia informacional de los usuarios no sea desarrollada, tanto las plataformas como los gobiernos deben contribuir en minimizar los efectos de la desinformación.

Palabras clave: Censura; Desinformación; MapReduce – Minería de textos.

Introduction

On October 5th, 2021, during a session in the United States Senate (*The New York Times*, 2021), it was pontificated that the U.S. government should control the content made available on social media in order to safeguard the health and safety of the public. The proposal was addressed by the testimony of Frances Haugen (*CBS News*, 2021), a data engineer (Wikipedia, 2023) from Facebook, currently Meta, who built her argument from her two-year work experience at the company, which would have allowed her to identify tools capable of modulating the information that travels in the digital ecosystem (*Folha de São Paulo*, 2021; *UOL*, 2021). This would allow curbing the dissipation of false news, limiting the dissemination of misleading narratives and even blocking access to a profile –a measure eventually taken against former US President Donald Trump by X (former Twitter) and Facebook (*BBC Brasil*, 2021)–. Coined as 'deplatforming,' this action was designed to remove someone from a platform (Oliveira, 2021).

Another figure in favor of regulating social media is Philippine journalist and 2010 Nobel Peace Prize winner Maria Ressa, who sees in social networks concrete threats to democracy (*Folha de São Paulo*, 2022), as they do not have an ethical commitment to an unbiased description of facts and truths. They set up trenches for polarizing discourses, which, in turn, dominate the narratives in virtual environments, in other words, social networks are the preferred medium for expressing the quintessence of extremist thought.

These discussions about content control on the Web were fostered from the recognition of the strength, scope, and speed of disinformation spreading

(deliberate attempts to confuse or manipulate people through transmission of dishonest information), incorrect information (inaccurate or incorrect content, disseminated without manipulative or malicious intent) and poor information (usage of data to discredit people and/or speeches) in the digital environment (Wardle and Derakhshan, 2018).
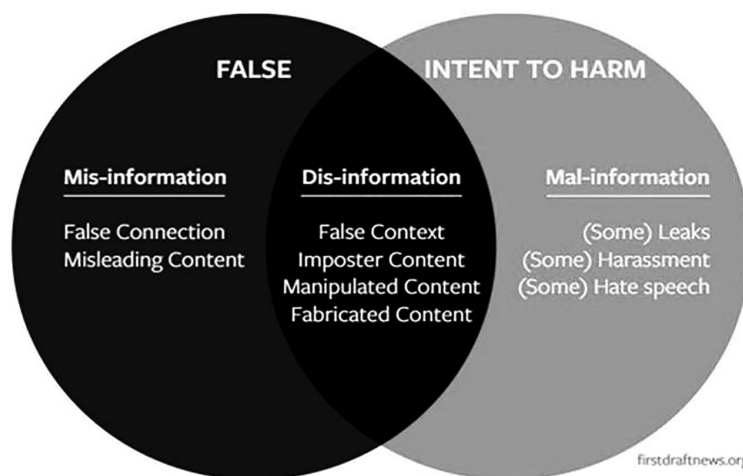

Figure 1. Mis-information, Dis-information and Mal-information
Source: Wardle and Derakhsahan, 2018: 48

It is important to note that censorship is not something recent, whether veiled or ostensible, it is something that permeates social relations. Customs, religion, political power and even the law (as in cases of child pornography, hate speech, national security, etc.) are invoked to justify its application. This "content control" (*Veja*, 2018) is exercised by several entities that serve as "ideological apparatus of the State" (Althusser, 1985), which are dedicated to filter the circulation of information, either by banning publications, or by extracting divergent discourse, or by repeating sponsored fake news.

Facebook, Instagram, X, and Twitch, among other content platforms, can censor information posted by users, through their conditions, policies, and terms of service, as happens daily with copyrighted material uploaded to said platforms (*Jornal do Comércio*, 2021). Lobbying from the intellectual property industry has succeeded in getting platforms to adopt mechanisms to cut live broadcasts or remove content from users when there are suspicions of copyright infringement.

Facebook and X have made evident, for example, their ability to censor the content on platforms, such as happened to former US President Donald Trump,

who was banned from X in 2021 and had his accounts blocked by Facebook in 2018 (*G1*, 2018). This filtering can be applied when publication contradicts the platform's rules, when compelled by judicial decisions, and when yielding to popular, economic or political appeals. Furthermore, recently, platforms that once adopted hands-off approach, with the development of algorithmic censorship could "shift towards *ex ante* forms of moderation; identifying and suppressing prohibited content as it is posted" (Cobbe, 2021).

From these findings, some inevitable questions arise: Is content control on these platforms adequate? Does the censorship exercised by platforms themselves hurt freedom of expression? What is the extent of the power given to platforms under the prerogative of content moderation? Can this faculty of moderation be confused with control of narratives? From this point on, the censorship of disinformation based on the principles of Information Science and Law, its application in digital environments, the role of regulators, as well as who should be responsible for the evaluation of content and eventual decision-making will be examined.

The goal is also to discuss the assertiveness of eventual control by the Judiciary, from the provocation of the interested party –The Brazilian Judiciary, by the Federal Constitution, cannot act *ex officio*; it is necessary that a party (individual or legal entity, an association, a federative entity, etc.) provoke it through a lawsuit–, to discuss searches in unstructured databases, with references to the fundamentals of textual mining, a brief presentation of MapReduce (a tool developed by Google Engineers to efficiently extract data and information from large repositories) and Bacen Jud (an application managed by the Central Bank of Brazil, chosen as an online tool model for linking the Judiciary and Web platforms).

In summary, in this article, disinformation and some of its presentation forms will be conceptualized, the effectiveness and risks of the use of content suppression and suspension/banning of profiles on the Web will be explored, the forms of censorship application (such as self-regulation and judicial control) will be described, the ways information search in large repositories will be addressed, and the need for the education of users will be evidenced.

FREEDOM OF SPEECH VERSUS CENSORSHIP

Freedom of expression is a fundamental human right recognized by the Universal Declaration of Human Rights, adopted by the United Nations General Assembly in 1948. Article 19 of this declaration states that: "Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive, and impart information and

ideas through any media and regardless of frontiers" (ONU, 1948). However, it is important to remember that many states members of the United Nations do not consider freedom of expression an absolute right, limiting it in certain circumstances, such as to protect national security or public health. Although these states advocate that limitations to freedom of expression must be justifiable, indispensable, and proportionate to prevent human rights violations, it is clear that they offer routes for abuse.

Most widespread idea is that censorship is directly opposed to the Principle of Freedom of Expression, defended in most advanced democracies, and textually inscribed in the Constitutions (the 'Law of Laws') of various countries. The Brazilian Constitution, for example, celebrates this principle in Articles 5 and 220, establishing:

Article 5

IV. the expression of thought is free, and anonymity is forbidden;

IX. the expression of intellectual, artistic, scientific, and communications activities is free, independently of censorship or license;

XIV. access to information is ensured to everyone and the confidentiality of the source shall be safeguarded, whenever necessary to the professional activity; [...]

Article 220. The manifestation of thought, the creation, the expression and the information, in any form, process or medium shall not be subject to any restriction, with due regard to the provisions of this Constitution.

§ 2° Any and all censorship of a political, ideological and artistic nature is forbidden. (*Constituição*, 1988)

This essential human right (freedom of expression) should not have any exception. Nonetheless, even the most progressist humanist thinkers tend to allow limiting its exercise when it is applied to block hate speech, the incitement to crime and violence, and in the extracting narratives that encourage xenophobia and discrimination against people on grounds of race, sex, skin color and sexual orientation.

One cannot, however, rely exclusively on the law to ensure respect for these limits (if admitted), as it (the law) is inconsistent. Time, social organization, beliefs, science, politics, scarcity, hunger, abundance, and many other factors constitute driving forces capable of changing the legal order and transforming the forbidden into the permitted, the illicit into the lawful. To support this assertion, it is enough to resort to recent history, more precisely, to the Holocaust perpetrated by the Nazis or to the 'Prohibition' imposed at the beginning of the last century in the United States of America. The first describes the genocide against

Jewish people under the approval of the propaganda and laws of the German state, the second exemplifies the unsuccessful attempt to ban alcoholic beverages sales in the US, which led to the strengthening of the mafia, the corruption of public agents, among other evils, until it was annulled.

In Brazil, during the most recent exception period –which historians establish as lasting from 1964 to 1985– censorship was official and ostensible, as granted by Law No. 5.250/1967. Although the state apparatus insisted that the censorship effort was abolished and dismantled by the Federal Constitution of 1988, the impetus to try to silence dissonant voices was never totally overcome in the country. The idea of media social control has returned to public debate in Brazil in the past few decades and, more recently, under the argument of the need for content moderation and information regulation on the Web (STF, 2020). This effort to control information traffic in the Web is materialized in Bills No. 2.630/2020 and 3.227/2021 and has been effectively applied in an ongoing judicial inquiry at the Supreme Court (2020), in which the demonetization of people and websites, the extraction of pages hosted on the Internet and the suspension of profiles and users were ordained (*Agência Brasil*, 2020).

The reactions to these proposals and the decision rendered at the Brazilian Supreme Court polarized the discussions because opinions differ as to the scope of these norms, the practical results and the risks generated to democracy, and the economy by the control and domination of the information flow. But there is no Brazilian exclusivity in this influx. The evidence in the world can be materialized in the testimony collected from the data engineer in the United States Senate, summarized in the initial paragraphs, in the banishment of the former American president from social networks and in the censorship applied to the Internet by the Chinese government (Ruan et al., 2016). The tipping point, therefore, is clear and needs to be addressed as there are social values at stake.

HUMAN BEHAVIOR AND MISINFORMATION ON THE INTERNET

Vosoughi, Roy and Aral (2018) show that misinformation spreads faster and more widely than truthful news on the Internet. People's fascination with the morbid, aberrations and the paroxysm of human actions –glimpses of the 'death drive' (Laplanche, 1991)–, exhibitionism, voyeurism, and the necessity of belonging, all these combined with the false security provided by distancing from physical reality are key factors that contribute to the success of this new world: the virtual environment. In this locus, people usually confuse distancing with anonymity; people forget, however, that they are under constant scrutiny, that they are being monitored, observed and that they constitute the most interesting

product of platforms, which produce and offer customized services to the user by selling their data and habits to goods and services producers.

Notably in the post-truth era, defined as "circumstances in which objective facts are less influential in the formation of public opinion than appeals to emotion and personal belief" (Ripoll and Canto, 2019), and although most people rationally favor limiting inappropriate content (real violence, hate speech, disinformation, among others), many will consume or share these same contents when hidden by the (illusory) distance provided by the screens of their computers and smartphones. Meaning that, even if society, as a whole, came together to interrupt or curb the flow of inappropriate content on digital platforms, many members would stop consuming this product if expectations were not met –expectations not confessed in public, usually linked to the contemplation and the stimulation caused by the exceptional, the extremes, the human folly–.

The moderation/regulation of social media content, in other words, the purging of undesirable content, at first, would not be enough to ensure that the digital environment became 'healthier,' 'civilized,' or 'inclusive.' It could, on the contrary, only provide the migration of the public to another environment that provides escape for untamed human urgencies. This reading reinforces the thought of science philosopher Mario Bunge (2004), who moves away from the classic scientific postulate of cause and effect, pointing out that there are many variables with the ability to influence an observable phenomenon. Human behavior, consumer relations and eventual political and religious polarizations are some of the elements that should be considered when examining the latest systems of censorship/moderation of content on the Internet.

In addition, the decision to censor/moderate content from social networks will necessarily go through a multifactorial decision, since both the contamination origin of the platform by false, exotic, or extreme content, and the attempt to asepsis of the social network suffer economic, social, political, religious, behavioral, and technical pressures. It is therefore imperative to decompose these forces and examine the systems and relationships between the components that will justify any of the decisions adopted by the State, the Judiciary or the operators who offer these services.

CENSORSHIP: A NECESSARY EVIL?

Freedom of expression should be examined in its two dimensions: the liberty to speak and the right to be informed. As a fundamental human right, freedom of expression is indispensable for the functioning of a free and just society (Habermas, 1997), and it must be protected against all forms of aggression. Nonetheless, even

the Inter-American Commission on Human Rights (IACHR) of The Organization of American States (OAS) has differentiated protected speech from that not protected by freedom of expression, meaning that freedom of expression must be cherished and emphasized when it comes to tolerate extreme criticism against public figures and agents of the states, but should be moderated "without prejudice to the presumption of coverage ab initio of all forms of human expression", when it comes to "propaganda for war and advocacy of hatred that constitute incitements to lawless violence", "direct and public incitement to genocide" or "child pornography" (IACHR, 2009).

This concern and limitation to freedom of expression expressed by the IACHR/OAS were due to the Internet, that potentially transformed a reader into a publisher, any viewer into a producer; the unidirectional mean of communication (e.g. radio, TV, newspaper) was surpassed by the Internet, that gave a voice to anyone in the world. Nevertheless, history proves that censorship accompanies social organizations since its cradle, being exercised by families, tutors, states, religions, ideologues, prelates, merchants, traditional medias and, currently, by great entrepreneurs of the digital age.

There would be ostensible censorship, that is easily identified, such as the banning of publications –for example, *The Satanic Verses* by Salman Rushdie was prohibited in several countries (*BBC Brasil,* 2012)– and the veiled form, sponsored by economic, religious and/or political interests. The latter is not always clear or perceived, especially because it is disguised with ideals, lies and various versions. Even disinformation, when is used as a tool to divert, to distance one's opinion from what is real or true, can be considered a form of censorship, especially because it would reverberate within the "networks of trust" (Bakir and McStay, 2017), "filter bubbles" (Pariser, 2012) and "echo chambers" (Posetti, 2018).

Internet, jet engines, space travel, communication, satellite navigation, the development of transport, medicine, the addition of many amenities to modern life, among other technologies developed in the last hundred years, are the fruit of the vicious war efforts or of the "balance of terror" (Aron, 1986). So, it is admissible to conceive that some good comes from evil, meaning that censorship, which, in appearance and form, diametrically opposes freedom and equality, could also have beneficial effects for the containment of "information disorders" (Wardle and Derakhshan, 2018) and the "hyperinformation" phenomenon (Moretzsohn, 2017) to spare people from hate speech and explicit violence.

Nonetheless, it is important to highlight the ethical implications of censorship for Information Science, since, as Guimarães, Pinho and Milani (2016) note, the Information Scientist cannot forget his ethical commitment, which obliges him or her to defend clarity, transparency, inclusion, guarantee of access, reliability and correctness of the information made available to individuals. In a nutshell, one

could postulate that even the act of denying access to inappropriate content should be explicitly announced by those obliged by the ethics code of Information Science.

THE ACTIONS OF THE STATE

The fight against disinformation should not justify the curtailment of freedom of expression; the healthy exchange of people and cultures should not be censored, and content moderation cannot become an instrument of political, religious, ideological, and economic control of people. It is noteworthy, therefore, that the expansion of disinformation studies and of the mechanisms of censorship/moderation of content in the World Wide Web is essential, especially because a human right is involved. Ideally, external censorship, understood as that exercised by another entity other than the users themselves, should not exist. Filters should be defined by the free and conscious individual, because "in the last resort, it is not the force of law but only the force of intelligence that can save a people from its own folly" (Schrader, 1993).

Although users themselves may have the skills to filter and differentiate fake from true content, if it is necessary to arbitrate eventual content regulation for platforms, as well as defining which rules are acceptable or not, it seems preferable to entrust this task to the Judiciary than to hand over said power to private entities, which base their actions on the stock market results (Seisdedos, 2021). The proposal of content moderation by the Judiciary draws its justification from the adage of 'choosing the lesser of two evils'. It could also be based on a skeptical reading of the current scenario, marked by war of narratives, information disorder, and the impossibility of refuting Michel Foucault's (2017) analyses of the truth, which would not exist outside of power.

But this option also carries risks, in the second half of 2022, news and social networks in Brazil turned away from the war between Ukraine and Russia and the effects of the pandemic caused by Covid-19 in order to broadcast an act of censorship perpetrated by the highest instance of Brazilian Justice, the Supreme Court: supported by both Law No. 12.965/2014 as well as recalcitrance of the platform, who had been systematically ignoring orders issued by the Judiciary, the Court decided to suspend the operation of the messaging application Telegram. Politicians *(Senado Notícias,* 2022) and several experts (Paiva, 2022) criticized the decision, pointing out that the targeted publications could be purged, but that the platform should be preserved, as dictated by the law that substantiated said decision. If there is a court order, the provider must promote the exclusion of content considered inappropriate. If no action is taken by the platform, then the proportional sanctions established by Article 12 of Law No.

12.965/2014 (known as the 'Internet Civil Framework'), should be applied. These range from warnings, through the application of fines and temporary suspensions, to the most extreme, the prohibition of activities.

In order to sanction this solution (the suppression of content, profiles and websites), the Brazilian Judiciary already has a ready model, namely, Bacen Jud, which consists of a digital platform "for communication between the Judiciary and financial institutions participating in the National Financial System, with technical intermediation of the Central Bank of Brazil" (BCB, 2018). With the implementation of Bacen Jud and the constitution of the Customer Registry of the National Financial System (CCS) (BCB, 2007), the Judiciary Branch itself was responsible for "the registration of orders in the system and the zeal for their compliance" (BCB, 2018: Art. 2, §1). The participating financial institutions remained "responsible for complying with court orders in the standardized form" (BCB, 2018: Art. 2, §2) while the Central Bank is responsible for the "operation and maintenance of the system" (BCB, 2018: Art. 2, §3).

Internet platforms and providers, in the same way as financial institutions, through a similar communication tool, technically modulated by another agency (for instance, the National Telecommunications Agency), would receive orders directly from Judges and/or duly qualified authorities to share and preserve data, extract content and suspend/cancel users, profiles and websites. The authority in question, therefore, would issue the order in the system, based on a standard form defined in a consortium by the participants involved, which would instantly be communicated to the managers of the services through which the censored content travels.

THE IDENTIFICATION OF UNDESIRABLE CONTENT

Although the idea, after all, represents nothing more than the replication of the Bacen Jud model and limits itself to giving enforceability to the letter of the law that provides the possibility of suppressing judicially censored content from the Web, many will vociferate against its implementation, among them, the same directors of large digital conglomerates that recently rose against judicial control (*G1*, 2021) of the content they host and claimed that without this external guardianship they would be able to more effectively, quickly and efficiently purge inappropriate content.

It is certain that one cannot overlook the fact that the Judiciary itself can be defiled by political influence (Baião et al., 2022), economic strength and financial power, which, as previously mentioned, happens in totalitarian states —China and North Korea impose political censorship on the Internet (Ruan et. al.

2016)–. One can, therefore, make use of techniques and tools of Text Mining in large repositories to seek greater effectiveness of the determined measures, since the speed of propagation of information on the internet makes it difficult to curtail the traffic of undesirable content, which, viewed on one platform, migrates to others in a brief interval, as in the process called 'viralization'.

Text mining

Text mining can be described as the fusion of data mining, machine learning (Mitchel, 1997) and natural language processing (Nadkarni, Ohno-Machado and Chapman, 2011). This technique is dedicated to overcoming the crisis resulting from information overload, supporting information retrieval and knowledge management (Feldman and Sanger, 2006). In any case, the first obstacle to overcome will certainly be research in large repositories.

To circumvent the difficulties arising from searching in these large repositories –an estimated 40 trillion gigabytes were generated in the world in 2020, which means that 9.1 thousand terabytes of data are generated every 6 minutes (*Exame*, 2021)– and to perform efficient searches in environments that offer structured and unstructured data, tools were developed by researchers who either created models or merged several solutions into new systems. Thus, there is MapReduce, a software developed by Google Engineers Jeffrey Dean and Sanjay Ghemawat, widely used for its efficiency status for processing large volumes of data and ease of use (Herodotou and Babu, 2011).

MapReduce is a processing model that allows operators to mine data (even those lacking large resources, as they could do it without data processing centers). Many computers –as many as available– can be interconnected to perform tasks in parallel (Dean and Ghemawat, 2004), and the failures are automatically treated, with the replacement of the task assigned to one computer transferred to another machine if the correct answer is not received, that is, the failure of one step does not paralyze processing or compromise the result. In sum, it can be said that MapReduce abbreviates large volumes of data into smaller sets, through the operations called MAP and REDUCE. In the first step, MAP, the program scans the database and separates them into pairs according to the stipulated 'Keys and Values'. In the next step, REDUCE, the results of the various machines are combined. Thus, "users specify the MAP function, which will process key/value pairs to generate an intermediate set of pairs, and the REDUCE function, which will perform the fusion of these results" (Dean and Ghemawat, 2004). These operations can be repeated countless times for the sake of research refinement.

To achieve a better understanding, let us take a practical and fictitious example of the tool's operation: voters of the defeated candidate (ECD), who received 46.5% of votes in the last election, express discontent with the result of the polls. Of this contingent of voters (46.5%), 23% openly declared their vote on social networks (ECD - RS), and 85% of these said they believed, through message sharing, that there was fraud in the elections. The operation of segregating voters and verifying how many defend the hypothesis of fraud represents the MAP operation, meaning, the program forces itself, through commands exposed in the algorithm, to 'break' the data and separate them according to categories (*Figure 2*).
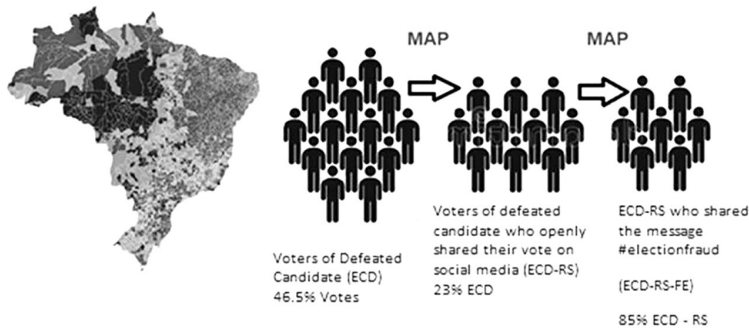

Figure 2. MAP function 'breaking' the data into groups
Source: Authors elaboration

In the next operation, REDUCE, this data is grouped, and new steps can be added. In the example, 70% of the ECD - RS - FE adhere, through social media, to the proposal to march on the capital to prevent the swearing-in of the victorious candidate.
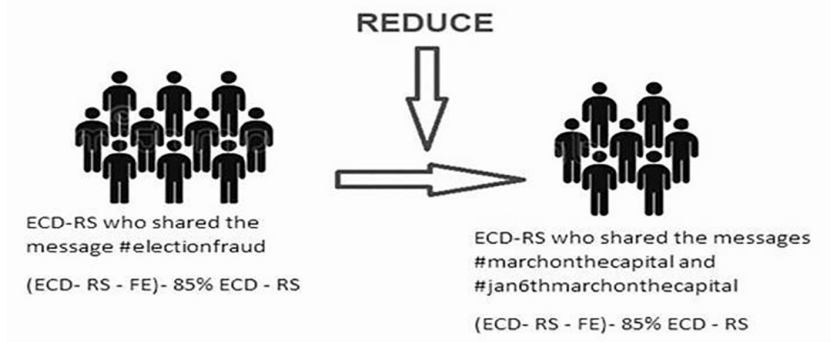

Figure 3. REDUCE function: grouping of results
Source: Authors elaboration

Refining the analysis, data related to the measurement of the flow of people in the days leading up to the alleged meeting in the capital could be added. In a new stage, therefore, new interpretable sets of data (information) would be generated, which would allow us to infer that the meeting of ECD in the capital is 'quite likely':

Notable increase of traffic in airports, tollbooths, buses, trains and rides heading to the capital between January 2nd and 5th

25% ECD-RS-FE-MC share messages documenting movement towards the capital (ECD-RS-FE-MC-DC)

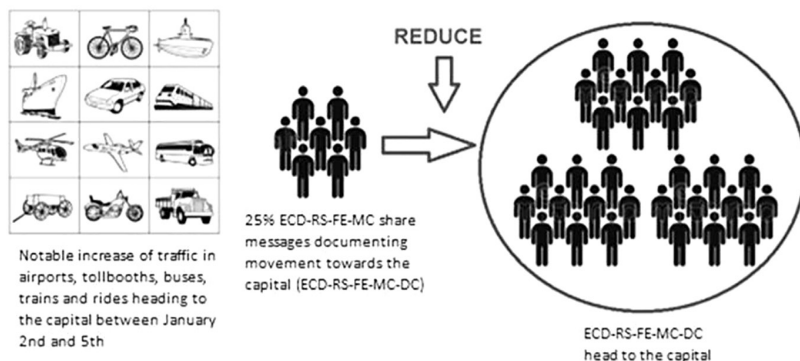ECD-RS-FE-MC-DC head to the capital

Figure 4. REDUCE function: groups the results and allows inferences
Source: Authors elaboration

Because its mechanisms are hidden (Dean and Ghemawat, 2004), the programmer has his tasks facilitated by simple commands. However, this does not mean that there is no work to do; on the contrary, the result of searches will depend exclusively on the intelligence of the 'business owner', who must know what exactly will be sought, the questions that should be answered, and the programmer who will design the algorithm.

This model can be used, as said, in repositories that contain unstructured data. Therefore, it performs searches in social networks and other sources based on natural language. For such a reach, the algorithm needs to be 'trained' or rather designed and prepared to correct or supplant semantic and syntactic inaccuracies. For example, the 'trained model' can scan social networks for patterns that identify misinformation, traffic generated by it, and the user responsible for producing and sharing said content.

Similar model has already been successfully implemented in a political environment, more precisely, during the re-election campaign of the President of the United States in 2012 (Hadoop Illuminated, 2023). With guaranteed access to users' data from social networks, the analysts and strategists of the victorious team did not invest in the advertising campaign model of the 1990s, characterized by commercial insertions in large media (especially on television), instead they focused on collecting and analyzing the data that traveled through social media, personalizing the ideological and economic ads and appeals, directing

them according to the profile and possible weaknesses, fears and prejudices of the voter/user, through social networks. There was developed customized content, tailored to manipulate the perception of reality and the behavior of the 'consumer' (voter).

Using available programs and systems, such as MapReduce, the Judiciary (or Government) could identify the sources and paths taken by undesirable content to determine its suspension. More than that, through trained algorithms, it would be possible to identify if there are bots or other malicious devices programmed to reverberate, circulate and disseminate disinformation *en masse*. MapReduce could therefore be the programming model used for searching large repositories, especially on the World Wide Web and social networks (Herodotou and Babu, 2011), full of unstructured data in natural language, and to respond to "an ecosystem that uses information disorder in its favor and actively provokes it" (Gitahy in Villen, 2020), especially in an environment permeated by "fake news", "post-truth", "deepfakes" and "alternative facts", which reveals "a scenario of hyperinformation'" (Moretzsohn, 2017) and offers the ideal camouflage for abundant portions of disinformation (Ripoll and Matos, 2020).

FINAL CONSIDERATIONS AND PERSPECTIVES

The world has been shrunk by electronic communication and more efficient means of transportation; borders are overcome with less effort and the exchange of ideas, goods, and the migration of people has become easier. This permeability, in turn, has brought new challenges, as people are exposed to a large amount of information; much of it, is true and useful; part of it is inadequate and harmless; another part is wrong and anodyne; and some, however, is false and carries with it the possibility of generating harmful repercussions, such as the invasion of the US Congress (Moraes and Nobre, 2022) and the Iraq War in 2003 (Hein, 2018).

The indisputable fact is that this movement, the flow of people, things and ideas and the clash of different worlds generates friction. However, the risks arising from this unprecedented cultural friction, although tangible, pale in the face of the achievements facilitated by the Internet in modern life as well as the promising results brought on by the communion of efforts in favor of progress, science, quality of life, as well as greater interaction between peoples.

Understanding and mitigating the possible deleterious effects of these advances, among them, disinformation through the Internet, are complex tasks because they do not only involve the use of computational tools to identify and extract inappropriate elements, but force one to deal with intricate components

of the human psyche, such as the desire for belonging and acceptance, imitation, behavioral modeling, "filter bubbles", "trusted networks" and "echo chambers", which stimulate (Bakir and McStay, 2017) the individual to share false content, premeditatedly or inconsequently, in the desire to see themselves among the herd.

Content censorship should not take precedence over freedom of expression, as blocking the flow of information has harmful effects not only on individuality, personal rights and guarantees, but also threatens social development and progress. Interventions to block undesirable content, if admitted, should be precise, based on legal precepts and should rely on advances in research related to natural language processing, machine learning and text mining, which can provide the development of tools suitable for the examination and coping with the spread of disinformation over the Internet. It should not be forgotten the importance of continuing to educate users with informational competences, in a way of preparing each and every one to select, collect, understand, and interpret the information that they gather, allowing them to develop and grow in this 'infodemic' context.

REFERENCES

Agência Brasil. 2020. "Senado aprova PL das Fake News; projeto segue para Câmara." June 30, 2020.
https://agenciabrasil.ebc.com.br/politica/noticia/2020-06/senado-aprova-pl-das-fake-news-projeto-segue-para-camara

Althusser, Louis. 1985. *Aparelhos ideológicos do Estado.* Rio de Janeiro: Graal.

Aron, Raymond. 1986. *Guerra e paz entre as nações.* 2nd ed. Brasília: Universidade de Brasília.

Baião, Bárbara, Fabio Graner, Fábio Zambeli and Flávia Maia. 2022. "Telegram é bloqueado no Brasil por decisão de Alexandre de Moraes". *Jota Info*, March 18, 2022.
https://www.jota.info/stf/do-supremo/alexandre-de-moraes-determina-o-bloque-io-do-telegram-no-brasil-18032022

Bakir, Vian, and Andrew Mcstay. 2017. "Fake News and The Economy of Emotions." *Digital Journalism* 6 (2): 154-75.
http://dx.doi.org/10.1080/21670811.2017.1345645

BBC Brasil. 2012. "'Versos Satânicos' não seriam publicados hoje, afirma Rushdie." September 17, 2012.
https://www.bbc.com/portuguese/noticias/2012/09/120917_salman_rushdie_entrevista_rw

BBC Brasil. 2021. "Trump é banido do Facebook e do Instagram por tempo indeterminado." January 7, 2021.
https://www.bbc.com/portuguese/brasil-55580191

BCB (Banco Central do Brasil). 2007. "Circular nº 3.347, de 11 de abril de 2007." Dispõe sobre a constituição, no Banco Central do Brasil, do Cadastro de Clientes do Sistema Financeiro Nacional (CCS).
https://www.bcb.gov.br/pre/normativos/circ/2007/pdf/circ_3347_v2_P.pdf

BCB (Banco Central do Brasil). 2018. "Regulamento Bacen Jud 2.0." April 2, 2018.
https://www.bcb.gov.br/Fis/pedjud/ftp/Regulamento-BACENJUD-02abr18.pdf

Brasil. 2003. *Lei nº 10.701/2003, de 09 de julho de 2003.* Altera e acrescenta dispositivos à Lei no 9.613, de 3 de março de 1998, que dispõe sobre os crimes de lavagem ou ocultação de bens, direitos e valores; a prevenção da utilização do sistema financeiro para os ilícitos previstos nesta Lei; cria o Conselho de Controle de Atividades Financeiras – Coaf, e dá outras providências.
https://www.planalto.gov.br/ccivil_03/leis/2003/l10.701.htm

Brasil. 2011. *Lei nº 12.527, de 18 de novembro de 2011.* Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências.
https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm

Brasil. 2012. *Decreto Lei nº 7724/2012, de 16 de maio de 2012.* Regulamenta a Lei nº 12.527, de 18 de novembro de 2011, que dispõe sobre o acesso a informações previsto no inciso XXXIII do caput do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição.
https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/decreto/d7724.htm

Brasil. 2014. *Lei nº 12.965/2014, de 23 de abril de 2014.* Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil.
https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm

Bunge, Mario. 2004. *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge.* Toronto: University of Toronto Press.

CBS News. 2021. "Facebook Whistleblower Frances Haugen Testifies before Congress." December 1, 2021. YouTube Video, 3:59:25.
https://www.youtube.com/watch?v=KCy4V2iMP6w

Cobbe, Jennifer. 2021. "Algorithmic Censorship by Social Platforms: Power and Resistance." *Philosophy & Technology* 34: 739-66.
https://doi.org/10.1007/s13347-020-00429-0

Constituição da República Federativa do Brasil. 1988.
http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm

Dean, Jeffrey, and Sanjay Ghemawat. 2004. "MapReduce: Simplified Data Processing on Large Clusters." Google, Inc.
https://static.googleusercontent.com/media/research.google.com/pt-BR//archive/mapreduce-osdi04.pdf

Exame. 2021. "Temos mais dados do que nunca. Como usá-los a nosso favor?" June 9, 2021.
https://exame.com/carreira/dados-uso-favor/

Feldman, Ronen, and James Sanger. 2006. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructered Data.* Cambridge: Cambridge University Press.

Folha de São Paulo. 2021. "Frances Haugen, a engenheira que quer 'salvar' o Facebook." October 5, 2021.
https://www1.folha.uol.com.br/tec/2021/10/frances-haugen-a-engenheira-que-quer-salvar-o-facebook.shtml

Folha de São Paulo. 2022. "Plataformas de internet estão destruindo a democracia, diz Nobel da Paz." June 28, 2022. YouTube Video, 10:52.
https://www.youtube.com/watch?v=ESjlTifAF_E

Foucault, Michel. 2017. *Microfísica do poder*. 5th ed. Rio de Janeiro: Paz e Terra.

G1. 2018. "Facebook bloqueia contas antes do início da eleição legislativa nos EUA." November 6, 2018.
https://g1.globo.com/economia/tecnologia/noticia/2018/11/06/contas-no-facebook-e-instagram-sao-bloqueadas-na-vespera-das-legislativas-nos-eua.ghtml

G1. 2021. "Executivos de Twitter, Google e Facebook são questionados sobre responsabilidade em invasão ao Congresso dos EUA." March 25, 2021.
https://g1.globo.com/economia/tecnologia/noticia/2021/03/25/executivos-do-twitter-google-e-facebook-falam-sobre-desinformacao-no-congresso-dos-eua.ghtml

Guimarães, José Augusto, Fabio Assis Pinho and Suellen Milani. 2016. "Theoretical Dialogs about Ethical Issues in Knowledge Organization: García Gutiérrez, Hudon, Beghtol, and Olson." *Knowledge Organization* 5 (43): 330-50.

Habermas, Jürgen. 1997. *Direito e democracia: entre facticidade e validade*. Translation by Flávio Beno Siebenneischler. Rio de Janeiro: Tempo Brasileiro.

Hadoop Iluminated. 2023. Open Source Hadoop Book.
https://hadoopilluminated.com/

Hein, Matthias von. 2018. "Guerra do Iraque, uma invenção americana." *Deutsche Welle*, April 9, 2018.
https://www.dw.com/pt-br/guerra-do-iraque-uma-iven%C3%A7%C3%A3o-americana/a-43309906

Herodotou, Herodotos, and Shivnath Babu. 2011. "Profiling, What-If Analysis, and Cost-Based Optimization of MapReduce Programs." *Proceedings of the VLDB Endowment* 4 (11): 1111-12.
https://doi.org/10.14778/3402707.3402746

IACHR (Inter-American Commission on Human Rights). 2009. *The Inter-American Legal Framework Regarding the Right to Freedom of Expression*. Accessed January 20, 2024.
https://shorturl.at/bnquB

Jornal do Comércio. 2021. "Redes sociais têm o direito de banir usuários por não seguirem regras." January 18, 2021.
https://www.jornaldocomercio.com/_conteudo/especiais/jornal_da_lei/2021/01/774424-redes-sociais-tem-o-direito-de-banir-usuarios-por-nao-seguirem-regras.html

Laplanche, Jean, and Jean-Bertrand Pontalis. 1991. *Vocabulário da psicanálise*. São Paulo: Martins Fontes.

Moraes, Tainá Felipe de, and Fábio Rodrigo Ferreira Nobre. 2022. "Mídias sociais e Fake News: desinformação como ferramenta antidemocrática no governo Trump." *Revista de Estudos Internacionais* 13 (1): 82-106.
http://dx.doi.org/10.29327/252935.13.1-5

Moretzsohn, Sylvia Debossan. 2017. "'Uma legião de imbecis': hiperinformação, alienação e o fetichismo da tecnologia libertária." *Liinc em Revista* 13 (2): 294-306.
https://doi.org/10.18617/liinc.v13i2.4088

Nadkarni, Prakash M., Lucila Ohno-Machado and Wendy W. Chapman. 2011. "Natural Language Processing: An Introduction." *Journal of The American Medical Informatics Association* 18 (5): 544-51.
https://doi.org/10.1136/amiajnl-2011-000464

Oliveira, Felipe. 2021. "Para especialistas, banir Trump cria precedentes sobre moderação nas redes." *Tilt UOL*, January 14, 2021.
https://www.uol.com.br/tilt/noticias/redacao/2021/01/14/banimento-de-trump-das-redes-sociais-ataca-liberdade-de-expressao-entenda.htm

ONU (Organização das Nações Unidas). 1948. *Declaração Universal dos Direitos Humanos*. https://www.ohchr.org/sites/default/files/UDHR/Documents/UDHR_Translations/por.pdf

Paiva, Letícia. 2022. "O que pensam os especialistas sobre o bloqueio do Telegram?" *Jota Info*, March 18, 2022. https://www.jota.info/coberturas-especiais/liberdade-de-expressao/stf-telegram-bloqueio-especialistas-18032022

Pariser, Eli. 2012. *The Filter Bubble*. New York: Penguin Random House.

Posetti, Julie. 2018. "News Industry Transformation: Digital Technology, Social Platforms and the Spread of Misinformation and Disinformation." In *Journalism, 'Fake News' and Disinformation: Handbook for Journalism Education and Training*, edited by Cherilyn Ireton and Julie Posetti, 55-69. Paris: United Nations Educational, Scientific and Cultural Organization.

Ripoll, Leonardo, and Fábio Lorensi do Canto. 2019. "Fake news e 'viralização': responsabilidade legal na disseminação de desinformação." *Revista Brasileira de Biblioteconomia e Documentação* 15 (1): 143-56. https://rbbd.febab.org.br/rbbd/article/view/1364.

Ripoll, Leonardo, and José Claudio Matos. 2020. "O contexto informacional contemporâneo: o crescimento da desinformação e suas manifestações no ambiente digital." *Informação@Profissões* 9 (1): 87-107. http://dx.doi.org/10.5433/2317-4390.2020v9n1p87

Ruan, Lotus, Jeffrey Knockel, Jason Q. Ng and Masashi Crete-Nishihata. 2016. "One App, Two Systems: How WeChat Uses One Censorship Policy in China and Another Internationally." *The Citizen Lab*, November 30, 2016. https://citizenlab.ca/2016/11/wechat-china-censorship-one-app-two-systems

Schrader, Alvin. 1993. *Fear of Words: Censorship and the Public Libraries of Canada*. Edmonton: University of Alberta Library.

Senado Notícias. 2022. "Lasier critica Alexandre de Moraes por bloqueio do Telegram." March 22, 2022. https://www12.senado.leg.br/noticias/noticias/materias/2022/03/22/lasier-critica-alexandre-de-moraes-por-bloqueio-do-telegram

Seisdedos, Iker. 2021. "A 'garganta profunda' do Facebook exige uma legislação para conter a rede social." *El País Brasil*, October 5, 2021. https://brasil.elpais.com/tecnologia/2021-10-05/a-garganta-profunda-do-facebook-exige-uma-legislacao-para-conter-a-rede-social.html

STF (Supremo Tribunal Federal). 2020. "Plenário conclui julgamento sobre validade do inquérito sobre fake news e ataques ao STF." June 18, 2020. https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=445860

The New York Times. 2021. "Whistle-Blower Unites Democrats and Republicans in Calling for Regulation of Facebook." October 5, 2021. https://www.nytimes.com/live/2021/10/05/technology/facebook-whistleblower-frances-haugen

UOL. 2021. "Ex-funcionária denuncia que Facebook engana repetidamente os usuários." October 5, 2021. https://economia.uol.com.br/noticias/efe/2021/10/06/ex-funcionaria-denuncia-que-facebook-engana-repetidamente-os-usuarios.htm

Veja. 2018. "Bloqueado na China, Facebook consegue licença para filial no país." July 25, 2018.
https://veja.abril.com.br/economia/bloqueado-na-china-facebook-consegue-licenca-para-filial-no-pais/https://veja.abril.com.br/economia/bloqueado-na-china-facebook-consegue-licenca-para-filial-no-pais/

Villen, Gabriela. 2020. "O ecossistema da desinformação." *UNICAMP Notícias*. August 20, 2020.
https://www.unicamp.br/unicamp/noticias/2020/08/20/o-ecossistema-da-desinformacao

Vosoughi, Soroush, Deb Roy and Sinan Aral. 2018. "The Spread of True and False News Online." *Science* 359 (6380): 1146-51.
http://dx.doi.org/10.1126/science.aap9559

Wardle, Claire, and Hossein Derakhshan. 2018. "Reflexão sobre a 'desordem da informação'": formatos da informação incorreta, desinformação e má informação. In *Jornalismo, Fake News and Desinformação: Manual para Educação e Treinamento em Jornalismo*, edited by Cherilyn Ireton and Julie Posetti, 46-58. Paris: United Nations Educational, Scientific and Cultural Organization.

Wikipedia. 2023. "Frances Haugen." Last modified April 11, 2023.
https://en.wikipedia.org/wiki/Frances_Haugen