

SAINT-PETERSBURG UNIVERSITY

Manuscript copyright

Kolotaev Yury Yurievich

**POST-TRUTH AND DIGITAL INFORMATION INFLUENCE IN
INTERNATIONAL RELATIONS: THEORETICAL CONCEPTUALIZATION
OF THREATS AND COUNTERMEASURES**

Scientific speciality 5.5.4. International Relations, Global and Regional Studies

Thesis

for a Candidate degree in political science

Translation from Russian

Scientific supervisor:

Doctor of historical sciences, professor
Khudoley Konstantin Konstantinovich

Saint Petersburg

2023

Contents

Introduction	3
Chapter 1: The notions of “Truth” and “Post-truth” in international security: A perspective on digital information influence and online manipulations	24
1.1 Truth, power and information influence through the lens of postpositivist and postmodernist epistemology	26
1.2 Post-truth and modern reflection on the role of truth and emotions in international relations	36
1.3 Critique and conceptual limitations of a scientific application of post-truth to the study of information influence in international relations	51
Chapter 2: The structure of digital information influence in international relations in the post-truth era	60
2.1. Methodology for studying the structure of information influence: individual and social aspects	61
2.2. Digital disinformation, hate speech, and the role of online platforms in international processes	73
Chapter 3: Information threats in the early 2020s	92
3.1. Artificial intelligence and the future prospects for digital information threats	93
3.2. Actors and models of countering information threats in international relations	102
Conclusion	121
Sources and Bibliography.....	127
Appendices	168

Introduction

Being strongly influenced by the digitalization process, the information and political space of modern states has found itself in an unprecedented situation in terms of the amount of information, messages, and data circulating in society. The growth in the volume of information is directly related to the ubiquity of the latest technologies that integrate the single information space of the Internet along with new network services and platforms. The redundancy of information in conjunction with the emergence of open social communication platforms (social networks) has become an important factor in the transformation of the political environment at various levels of the system of international relations. This transformation is associated with an increase in the publicity of political processes and the openness of information flows, the formation of two-way communication between society and government. At the same time, the digital information environment, with a clear trend towards democratization, is under a direct threat of information influence.

Digital information influence reflects the process of influencing a person individually or society as a whole through intentionally or unintentionally disseminated information in the digital space. This kind of impact, depending on the initial purpose, can be associated with both a positive and a negative or manipulative agenda. In the third decade of the 21st century, characterized by the growing polarization of the global and national information space, the problem of information manipulation in world politics becomes especially dangerous. An important condition that stimulates the process of polarization is the politics or condition of “Post-truth” reflecting the deformation of the public consensus about the significance of facts and verified information through the lens of an emotional assessment of reality. Post-truth has led to the actualization of disinformation (“Fakes”) and hate speech in the new digital environment. The immense amount and intensity of information flows transform the users’ and politicians’ ability to separate true and false information resulting in emotional or stereotyped thinking that determines the response to information stimuli.

Such social and political changes became prerequisites for the use of digital information influence while implementing political goals. In a situation of growing global instability in international relations, this carries the danger of inciting conflicts and growing hostility through the dissemination of politically motivated disinforming content. Conflicting actors inevitably turn to new forms of manipulation to disseminate information that can turn society against or set to support a particular political group.

The combination of these factors forms a broad problem area that needs to be studied and theoretically conceptualized due to the growing amount of unsystematized empirical data. Much of the research in the field of information influence¹ has focused on the study of topical linguistic articulation, public problematization, and the identification of precedents for the spread of digital misinformation. At the same time, the diversity of research in the field of studying the practical side of information manipulation in the realities of the 21st century requires a systematic political science theoretical analysis because, without proper scientific and methodological elaboration, the problem of information influence remains an issue open to socio-political speculation. A consistent analysis through the prism of the latest concepts of international relations theory can help to eliminate these costs, which is the main priority of this dissertation.

Today, under growing confrontation in the information and communication spheres, Russian science urgently needs to develop new theoretical approaches and methodological frameworks independent of external politicized influence. Strengthening the Russian scientific school of systemic analysis of digital information influence in all its diversity is an important factor both in the formation of its national approach to

¹ Research Anthology on Fake News, Political Warfare, and Combatting the Spread of Misinformation / Information Resources Management Association. – IGI Global, 2020. – 653 p.; Mustonen-Ollila E. Information Influence in Society's Information Environment / E. Mustonen-Ollila, M. J. Lehto, J. Heikkonen // *Journal of Information Warfare*. – 2020. – Vol. 19. – №. 4. – P. 70-88; Hammond-Errey M. Understanding and assessing information influence and foreign interference / M. Hammond-Errey // *Journal of Information Warfare*. – 2019. – Vol. 18. – №. 1. – P. 1-22; Huskaj G. State-of-the-art of Scientific Research on Disinformation / G. Huskaj, S. A. Axelsson // *European Conference on Cyber Warfare and Security*. – 2023. – Vol. 22. – №. 1. – P. 602-609; Rustamova L.R. Information Influence on the Age of "Post-Truth" and Fake-News / L.R. Rustamova, B.A. Barabash // *Political Science Issues*. – 2018. – Vol. 8. – №. 5. – P. 23-30. (In Russian); Petrishev E.V. Destructive informational influence in the context of countermeasures of challenges and threats of national security in modern Russia/ E.V. Petrishev // *Central Russian Journal of Social Sciences*. – 2017. – Vol. 12. – №. 4. – P. 83-90. (In Russian); Prikhodko M.V. Linguistic model of information war: structural elements and impact levels / M.V. Prikhodko // *Bulletin of the Moscow Region State University. Series: Linguistics*. – 2023. – №. 3. – P. 57-71. (In Russian).

responding to new digital threats and in the active promotion of an objective image of the country, free from external and internal disinformation or online manipulation influence.

The lack of theoretical development in the 21st-century information influence problems, including online disinformation and post-truth politics, was also noted at the key Russian research conventions in international relations. Thus, at the Congress of the Russian Association of International Studies (12-14 October), as well as at the conference “Digital International Relations” (25-26 October), held at the Moscow State Institute of International Relations (University) of the Ministry of Foreign Affairs of the Russian Federation (MGIMO University), researchers and practitioners noted a rising demand for domestic research on the theoretical elaboration of digital manipulation at the level of analysis of the main actors and variables of the process of information influence. Solving this scientific problem is an absolute priority in the study of digital threats.

This lack of theoretical conceptualization, together with the formation of the conditions of the global information crisis of the early 2020s, associated with the growing conflict between the West and Russia, deteriorating bilateral relations between various regional powers, increasing examples of military confrontations, as well as several non-political threats, forms a **high relevance of the study** of the problems of information manipulation in the realities of the digital society, with a particular emphasis on the issue’s theoretical dimension. Global information challenges overlap with various examples of information campaigns at the national level, resulting in the increasing relevance of “Disinformation” in the mid-2010s. Today’s international conflicts need a better understanding not only of the political, military, or social potential of the parties but also their ability to actively influence the information environment of the opponent, his representation in the digital space. The level of the information influence techniques as well as the ability to apply advanced digital technologies to spread a strategic narrative indicate this ability. These issues, with all the plurality of available examples of information influence, create a need for the research community to theoretically and methodologically model information influence as an object of research in the political and international space.

Additionally, the need for the universal theorization of the international information influence processes appears due to currently unresolved problems limiting empirically oriented research. In particular, direct empirical evidence showing a correlation between fakes and real manipulations of the political process is still lacking. This aspect remains largely debatable, because causal relationships may be based on spurious correlation or be influenced by a certain political consensus. However, the very fact of the disinformation's presence in the information space and its securitization through the development of countermeasures remains important.

To better counter the mechanisms of information manipulation, meaningful and conceptual studies of specific examples of threats expressed in disinformation and hate speech are required. Understanding the main actors, proliferation stages, and social and psychological influence tools is more than ever in demand in an information society. At the same time, it is important to study the information threats more deeply, by going beyond the purely political sphere. World problems and cataclysms, such as the COVID-19 pandemic, illustrate that the focus on disinformation and other problems of the post-truth era only from a political perspective distorts the perception of existing threats. The active involvement of users in digital communications makes ordinary people also participants in digital manipulations while producing them consciously or unconsciously. In exceptional cases, this gives rise to involuntary avalanches of false content on the web, causing a digital infodemic, i.e. an uncontrolled viral spread of harmful information.

This complexity of the problem of digital information influence, extending to world politics, requires a multifaceted analysis to avoid the reductionism of digital threats. Such an approach involves a consistent conceptualization of each of the main concepts of information influence, establishing an ontological link between them and extrapolating them to the system of international relations. A key component is the inclusion of security and actorship issues in studying the activities of information influence actors and intermediaries. This approach makes it possible to address the study of threats from the point of view of social process adaptation and the development of collective and global resilience to disinformation. In this regard, it is particularly important to theoretically model the environment of information influence and the content of threats associated with

negative information influence of any form by identifying psychological, informational, and social triggers implemented in international relations.

These priorities of scientific research in the field of information influence correspond to such **areas of international relations studies** as the development of theory and methodology, methods of analysis, conceptual apparatus, as well as the research of information, cognitive, bio, and other new technologies in world politics. The inclusion of post-truth politics and the emergence of new actors in the process of international information influence in the research agenda also reflects the content of such areas as megatrends, international security, the activities of non-state actors, as well as subjective factors in the study of international relations.

A supporting component in the study of information influence is the analysis of initiatives representing a collective response to post-truth, as reflected in the actions of various national, regional, and supranational practices. Many global actors, including the United Nations (UN) and the United Nations Educational, Scientific and Cultural Organisation (UNESCO), as well as regional actors such as the Commonwealth of Independent States (CIS) and the Collective Security Treaty Organisation (CSTO), have engaged in this process. Similar trends can be observed at the national level and even in the private sector through transnational corporations owning digital platforms. All these circumstances require consolidation of the available management practices and a reflection on the prospects for developing a multilateral approach capable of activating all of the systems resources to efficiently transform the international social and political space, which constantly faces information threats.

Drawing on the identified issues, **the aim of this dissertation** is to create a comprehensive theoretical and methodological model for the analysis of digital information influence in international relations and world politics, relying on the consistent consideration of conceptual and practical aspects of the current state of digital and information uncertainty. This aim is based on a systematic study of factors shaping the behavior of various actors both at the stage of information manipulations and at the stage of counteracting them.

Achieving the aim of the dissertation requires solving the following **tasks**:

- 1) Study the modern philosophical and scientific understanding of “Truth” and “Falsehood” in the context of poststructuralist and postmodern studies;
- 2) Establish a conceptual framework for “Post-truth” to describe the conditions of the modern communication environment;
- 3) Determine the structural relationship between the main actors (subject, mediator, and object) of the information influence process;
- 4) Provide a theoretical characterization of the key information manipulation threats in the post-truth era, pointing out the specifics of disinformation and hate speech as independent categories;
- 5) Identify potential directions for the evolution of digital and information threats;
- 6) Identify the existing range of information influence counteraction practices at the global, national, and local levels through the lens of “Resilience theory” and “Multistakeholderism” as a universal management logic;
- 7) Develop essential scientifically relevant characteristics of “resilience” and “multistakeholderism” as relevant management concepts in countering disinformation.

The research object is the digital information influence as a process in an information society of the 21st century. At the same time, **the research subject** consists of theoretical ideas about specific information threats - disinformation and hate speech - in the context of their modern scientific conceptualization, as well as current and promising models and methods of countering them.

The study relies on a combination of several **theoretical approaches**. A poststructuralist understanding of “Truth”, based on the works of J.F. Lyotard, and M. Foucault, as well as the postmodern insight into the information environment by J. Baudrillard, is at the core of the research. The relationship between “Truth” and “Power”, “Communication” and “Consensus” receive particular attention since they become the main categories of a theoretical understanding of information influence in the context of world politics. The “Post-truth” concept represents the contemporary interpretation of the dichotomy of “Truth” and “Falsehood”, which is pivotal in matters of strategic

manipulation and disinformation. The study develops the scientific component of “Post-truth”, providing new opportunities for analyzing the information factor in international relations despite the current gaps and shortcomings of the concept.

A special place in the analysis of the modern digital environment belongs to the emotional and cognitive aspects of information and users' actions since the growing emotionality of the social discourse constitutes a vital indicator of the post-truth era. To this end, the study addresses the tackling of information threats through bounded rationality theory at the individual level and the “Emotional regimes” concept at the social level. And among other important theoretical approaches, the research uses the concepts of “Normativity”, “Resilience” and “Multistakeholderism”. The latter two represent not only separate theoretical concepts but also an integral part of the research methodology.

The research methodology includes the use of general scientific and special methods. Among the general scientific methods, the study relies on:

- comparative analysis used to compare both theoretical approaches to post-truth and disinformation, and to study the existing practices of various political and social institutions in tackling information threats;

- the dialectical method deconstructs various post-truth politics dichotomies and binary oppositions, e.g. “Truth” and “Falsehood”, as well as “Factual truth” and “Emotions”. Through dialectics, information influence constitutes a situation, in which the subject, mediator, and object primarily in opposite states, become merged through the unifying nature of the information space. The behavior models of each actor are seen as a simultaneous unity and rivalry of the social and individual experience;

- classification allows the building of hierarchical relationships between key concepts and their attributes.

Among the special political science methods, the study relies on:

- system analysis enabling the study of the main actors (state, non-state, interstate actors), their structural and functional connection, and of the system levels (global, macro-regional, national, and local levels) where disinformation is tackled;

- the institutional method allows to analyze formal and informal (norms, ideology) governance institutions within individual actors, especially the EU;

- event analysis is used to sequence processes and events (their causes and consequences) that determine the development of both theoretical concepts (“Post-truth”) and the actor’s actions. The latter is especially important in studying the evolution of EU policy from 2014 to 2022 ;

- discourse analysis reveals communication patterns and strategic narratives characterizing the EU’s approach to digital challenges and the securitization of threats in the social discourse;

- agent modeling structuring the incentives and reactions of different actors to information threats and modeling the perception of manipulated information through the lens of the subject, the mediator, and the object of information influence.

The implementation of these methods relied on several groups of **data and information sources**:

The first group of sources includes normative and legal documents that establish the rules, principles, and laws in the digital sphere. They can be categorized both thematically, including legal acts and laws on disinformation² and hate speech³, and institutionally, depending on which actor adopted the relevant law: states⁴ or global institutions (e.g., UN recommendations)⁵.

The second group of sources contains terms of service⁶ or community guidelines⁷ which serve as the regulatory basis for self-regulation and moderation on online platforms and services.

² LOI n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l’information (1) / NOR : MICX1808389L, JORF n°0297 du 23 décembre 2018. – 2018; LOI n° 2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur internet (1) / NOR : JUSX1913052L, JORF n°0156 du 25 juin 2020. – 2020.

³ Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz - NetzDG) vom 1. September 2017 / Bundesgesetzblatt, T. I Nr. 61, S. 3352. – 2017.

⁴ Op. cit. (LOI n° 2018-1202 du 22 décembre 2018...); Op. cit. (Gesetz zur Verbesserung der Rechtsdurchsetzung...).

⁵ The UN Strategy and Plan of Action / United Nations [Electronic resource]. – URL: <https://www.un.org/en/hate-speech/un-strategy-and-plan-of-action-on-hate-speech> (accessed: 04.12.2022); United Nations. Resolution adopted by the General Assembly on 24 December 2021 A/RES/76/227 / United Nations – 2022. – P. 1-4.

⁶ Terms of Service / Telegram [Electronic resource]. – URL: <https://telegram.org/tos/terms-of-service-for-telegram-premium> (accessed: 21.02.2023); YouTube. Hate speech policy - YouTube Help [Electronic resource]. – URL: <https://support.google.com/youtube/answer/2801939?hl=en> (accessed: 10.11.2021).

⁷ YouTube Community Guidelines - How YouTube Works (in Russian) / YouTube [Electronic resource]. – URL: <https://www.youtube.com/howyoutubeworks/policies/community-guidelines/#community-guidelines> (accessed: 21.02.2023); Community Guidelines (in Russian) / Tiktok [Electronic resource]. – URL: <https://www.tiktok.com/community-guidelines> (accessed: 20.12.2022).

Statistical data form the third group of sources, including information and data from bibliographic and citation databases, in particular Scopus, used for bibliometric analyses of individual concepts under study.

The fourth group of sources is news articles published by various news agencies⁸. They serve as the basis of event analysis, including Russian and foreign publications.

The fifth group of sources covers a wide range of electronic resources reflecting the information about the activity of various public organizations. These include websites of fact-checking communities⁹, media literacy promotion organizations, etc. The electronic resources help to analyze the self-representation of various stakeholders from the civil society.

Turning to the degree of **scientific development of the topic**, it should be emphasized that the problem of information influence has a multifaceted representation in the scientific literature since it includes various areas of research related to propaganda, disinformation, post-truth, hate speech, as well as modern digital technologies and online platforms.

The most fundamental aspects of information manipulation are reflected in the classic studies of propaganda and manipulation by H.D. Lasswell¹⁰, R. Braddock¹¹, and J. Ellul¹². These studies offer examples of interaction models of various communicators, paying attention to the motives and consequences. M. McLuhan's concept on the role of media adds important insights into communicative practices¹³, stressing the special role of the mediator in shaping the context and meaning of any information. There are other

⁸ British public still believe Vote Leave '£350m a week to EU' myth from Brexit referendum / The Independent [Electronic resource]. – URL: <https://www.independent.co.uk/news/uk/politics/vote-leave-brexit-lies-eu-pay-money-remain-poll-boris-johnson-a8603646.html> (accessed: 08.02.2023); Freedland J. Post-truth politicians such as Donald Trump and Boris Johnson are no joke | Jonathan Freedland [Electronic resource]. – URL: <http://www.theguardian.com/commentisfree/2016/may/13/boris-johnson-donald-trump-post-truth-politician> (accessed: 24.01.2021).

⁹ CaptainFact.io [Electronic resource]. – URL: <https://captainfact.io/> (accessed: 22.12.2022).; Faktisk. [Electronic resource]. – URL: <https://www.faktisk.no/> (accessed: 22.12.2022).

¹⁰ Lasswell H.D. Propaganda technique in the World War / H.D. Lasswell. – New York: P. Smith, 1938; Lasswell H.D. The Theory of Political Propaganda / H.D. Lasswell // The American Political Science Review. – 1927. – Vol. 21. – № 3. – P. 627-631.

¹¹ Braddock R. An Extension of the "Lasswell Formula" / R. Braddock // Journal of Communication. – 1958. – Vol. 8. – № 2. – P. 88-93.

¹² Ellul J. Propaganda: The Formation of Men's Attitudes. Propaganda / J. Ellul; transl. K. Kellen, J. Lerner. – 1st edition. – New York: Vintage, 1973. – 352 p.

¹³ McLuhan M. Understanding Media: The Extensions of Man. Understanding Media / M. McLuhan, L.H. Lapham. – Reprint edition. – Cambridge, Mass: The MIT Press, 1994. – 389 p.

studies focused on media influence and “Media effects”¹⁴. i.e., media influence on collective opinion. Numerous important contributions in the field of information impact modeling belong to C. Kopp¹⁵, who created an information-theoretic model of information wars, which was later introduced into the analysis of disinformation.

Psychology and cognitive research are also of great importance in information influence research, especially the concept of “Bounded rationality” developed by H.A. Simon¹⁶ and D. Kahneman¹⁷. Despite being aimed at studying fundamental cognitive biases and heuristics, the studies themselves help to explain the manipulative phenomena in the information space of the digital age¹⁸. Similarly, works on emotionality and collective sentiments¹⁹, defined, in particular, in terms of “Emotional regimes” (developed by W.M. Reddy)²⁰, are also important in the context of applying psychology to information influence research.

The modern vision of digital information threats, expressed in the scientific and public discourse, pays special attention to three key concepts: “Post-truth”,

¹⁴ Scheufele D.A. Framing, agenda setting, and priming: The evolution of three media effects models / D.A. Scheufele, D. Tewksbury // *Journal of Communication*. – 2007. – Vol. 57. – P. 9-20; Scheufele D. Framing As a Theory of Media Effects / D. Scheufele // *The Journal of Communication*. – 1999. – Vol. 49. – P. 103-122; Pan Z. Priming and Media Impact on the Evaluations of the President’s Performance / Z. Pan, G.M. Kosicki // *Communication Research*. – 1997. – Vol. 24. – № 1. – P. 3-30; Brewer P.R. Priming or Framing: Media Influence on Attitudes Toward Foreign Countries / P.R. Brewer, J. Graf, L. Willnat // *Gazette (Leiden, Netherlands)*. – 2003. – Vol. 65. – № 6. – P. 493-508.

¹⁵ Kopp C. Classical deception techniques and perception management vs. the four strategies of information warfare / C. Kopp // *Conference Proceedings of the 6th Australian Information Warfare and Security Conference*. – School of Information Systems, Deakin University, 2005. – P. 81-89; Kopp C. Shannon, Hypergames and Information Warfare / C. Kopp // *Journal of Information Warfare*. – 2003. – Vol. 2. – № 2. – P. 108-118; Kopp C. Information-theoretic models of deception: Modelling cooperation and diffusion in populations exposed to “fake news” / C. Kopp, K.B. Korb, B.I. Mills // *PLOS ONE*. – 2018. – Vol. 13. – № 11.

¹⁶ Simon H.A. Human Nature in Politics: The Dialogue of Psychology with Political Science / H.A. Simon // *The American Political Science Review*. – 1985. – Vol. 79. – № 2. – P. 293-304.

¹⁷ Kahneman D. On the reality of cognitive illusions / D. Kahneman, A. Tversky // *Psychological Review*. – 1996. – Vol. 103. – P. 582-591; Tversky A. Judgment under Uncertainty: Heuristics and Biases / A. Tversky, D. Kahneman // *Science*. – 1974. – Vol. 185. – № 4157. – P. 1124-1131; Kahneman D. *Thinking, Fast and Slow* / D. Kahneman. – London: Macmillan, 2011. – 512 p.

¹⁸ Falling for fake news: the role of political bias and cognitive ability / E.C. Tandoc [et al.] // *Asian Journal of Communication*. – 2021. – Vol. 31. – № 4. – P. 237-253; Information overload for (bounded) rational agents / E.M. Pothos [et al.] // *Proceedings of the Royal Society B: Biological Sciences*. – 2021. – Vol. 288. – № 1944.

¹⁹ Bakir V. Empathic Media, Emotional AI, and the Optimization of Disinformation / V. Bakir, A. McStay // *Affective Politics of Digital Media*. – Routledge, 2020; Bakir V. Fake News and The Economy of Emotions / V. Bakir, A. McStay // *Digital Journalism*. – 2018. – Vol. 6. – № 2. – P. 154-175; Wahl-Jorgensen K. Media coverage of shifting emotional regimes: Donald Trump’s angry populism / K. Wahl-Jorgensen // *Media, Culture and Society*. – 2018. – Vol. 40. – № 5. – P. 766-778; Wahl-Jorgensen K. *Emotions, media and politics: Contemporary political communication* / K. Wahl-Jorgensen. – Cambridge, UK: Polity, 2019; Martyanov D.S. Emotional public sphere: Polarization of paralinguistic internet discourse / D.S. Martyanov, G.V. Lukyanova // *Vestnik Moskovskogo universiteta. Seriya 10. Zhurnalistika*. – 2021. – № 2. – P. 25-48. (In Russian).

²⁰ Reddy W.M. *The Navigation of Feeling: A Framework for the History of Emotions*. The Navigation of Feeling / W.M. Reddy. – Cambridge: Cambridge University Press, 2001; McGrath L.S. *Historiography, affect, and the neurosciences* / L.S. McGrath // *History of Psychology*. – 2017. – Vol. 20. – № 2. – P. 129-147.

“Disinformation” (or “Fake news”), and “Hate speech”. Among them, “Post-truth” is the broadest and most discussed concept, incorporating all the other manifestations of information manipulations. Its theoretical development has several stages of conceptualization. The first mentions of this concept by such authors as S. Tesich²¹ and R. Keyes²², who shaped the theoretical problem field itself, can be attributed to an earlier stage. The next step is the actualization of the concept during the 2016 events, which have become central examples and starting points for a new phase of post-truth research. These events include the 2016 referendum on Britain's withdrawal from the European Union (EU) and a range of alleged attempts to influence electoral processes, including the most prominent example of the 2016 election of Donald Trump as US President. Each of these events has led to several essential monographs on the philosophical and social understanding of the phenomenon (e.g., D. Ball²³, M. d’Ancon²⁴, E. Davies²⁵, and S. Fuller²⁶). These works have become a catalyst for scientific discussion about the relevance of the “new” concept and phenomenon.

Subsequently, the interest in post-truth rose, which was reflected both in the integration of the concept into multiple scientific disciplines²⁷, and in the translation of the English concept into other languages. To date, post-truth gradually becomes an independent subject of study in the context of political science, i.e., interpreted in terms of power²⁸. An active increase in the number of publications on post-truth also happens

²¹ Tesich S. A government of lies. / S. Tesich // *Nation*. – 1992. – Vol. 254. – № 1. – P. 12-14.

²² Keyes R. The post-truth era: dishonesty and deception in contemporary life. *The post-truth era* / R. Keyes. – 1st ed. – New York: St. Martin’s Press, 2004. – 312 p.

²³ Ball J. Post-truth: how bullshit conquered the world. *Post-truth* / J. Ball. – London: Biteback Publishing, 2017. – 306 p.

²⁴ D’Ancona M. Post truth: the new war on truth and how to fight back. *Post truth* / M. D’Ancona. – London: Ebury Press, 2017. – 167 p.

²⁵ Davis E. Post-truth: why we have reached peak bullshit and what we can do about it. *Post-truth* / E. Davis. – London: Little, Brown, 2017. – 347 p.

²⁶ Fuller S. The Post-Truth About Philosophy and Rhetoric / S. Fuller // *Philosophy & Rhetoric*. – 2017. – Vol. 50. – № 4. – P. 473-482; Fuller S. Post-Truth: Knowledge As A Power Game. *Post-Truth* / S. Fuller. – New York: Anthem Press, 2018. – 218 p.

²⁷ Effron D.A. The moral psychology of misinformation: Why we excuse dishonesty in a post-truth world / D.A. Effron, B.A. Helgason // *Current Opinion in Psychology*. – 2022. – Vol. 47. – P. 101375; Post-Truth, Philosophy and Law / ed. A. Condello, T. Andina. – 1st edition. – Routledge, 2020. – 172 p.; Gunkel D.J. The medium of truth: media studies in the post-truth era / D.J. Gunkel // *Review of Communication*. – 2019. – Vol. 19. – № 4. – P. 309-323; Heinrich S. Medical science faces the post-truth era: a plea for the grassroots values of science / S. Heinrich // *Current Opinion in Anaesthesiology*. – 2020. – Vol. 33. – № 2. – P. 198-202.

²⁸ Kalpokas I. A Political Theory of Post-Truth / I. Kalpokas. – Cham: Springer International Publishing, 2019. – 135 p.; Block D. Post-Truth and Political Discourse / D. Block. – Cham: Springer International Publishing, 2019. – 130 p.; Giusti S. Democracy and Fake News: Information Manipulation and Post-Truth Politics. *Democracy and Fake News* / S. Giusti, E. Piras. – Routledge, 2021. – 246 p.

in Russian science. The works of the author teams of the St. Petersburg State University²⁹, as well as such authors as N.F. Ponomarev³⁰, I.V. Ponkin³¹, and S.V. Chugrov³² develop the concepts in Russia. They make a special emphasis on post-truth as an indicator of the decline and erosion of Western institutions³³, resulting from growing mistrust and mutual misunderstanding between society and government in the 21st century³⁴. However, “Post-truth” continues to develop today beyond Western political systems, as evidenced by works from the Middle East³⁵ and from the “Global south” generally³⁶.

In addition to post-truth, modern research on information manipulation refers to fake news and disinformation as specific forms or tools of influence. Most of the research in this area focuses either on the classification of threats or on the development of countermeasures. Among the most significant are the classifications proposed by S.N. Ilchenko³⁷, E.R. Romanova³⁸, E. Kapantai³⁹, and S. Lewandowski⁴⁰. Similarly notable

²⁹ “Post-Truth Politics” and Populism / ed. O.V. Popova. - St. Petersburg: OOO "Skifia-print", 2018. - 216 p. (In Russian); Post-truth politics in the modern world / ed. O.V. Popova. – Proceedings of the Russian scientific conference with international participation "Post-truth politics and populism in the modern world" September 22–23, 2017. - St. Petersburg: LLC "Skifia-print", 2017. - 282 p. (In Russian); Governability and discourse of virtual communities in the conditions of post-truth politics / I.A. Bykov [et al.]; ed. D.S. Martyanov. - St. Petersburg: ElekSys, 2019. (In Russian).

³⁰ Ponomarev N.F. Postmodern strategic communications. Post-truth, memes, transmedia / N.F. Ponomarev. – «KnoRus» Publishing House, 2023. - 174 p. (In Russian).

³¹ Ponkin I.V. Non-classical wars / I.V. Ponkin. – “INFRA-M”, 2019. (In Russian).

³² Chugrov S.V. Post-Truth: Transformation of Political Reality or Self-Destruction of Liberal Democracy? / S.V. Chugrov // Polis. Political Studies. – 2017. – № 2. – P. 42-59. (In Russian).

³³ Denilkhanov A.K. Post-truth and the liberal paradigm of postmodernity / A.K. Denilkhanov // Bulletin of Moscow University. Series 12. – 2021. – № 6. – P. 42. (In Russian).

³⁴ Eremina N.V. “New Normal”: The End or the Beginning? / N.V. Eremina // The New Past. – 2021. – № 3. – P. 198–207; Tianyang F. Socio-political turbulence and its communicative consequences / F. Tianyang, N.Yu. Markushina // Law and Power. – 2023. – № 2. – P. 4-6. (In Russian).

³⁵ Cosentino G. Post-truth politics in the Middle East: the case studies of Syria and Turkey / G. Cosentino, B. Alikasifoglu // Artnodes. – 2019. – № 24. – P. 91-100; Medeiros J. de. Conspiracy Theory in Turkey: Politics and Protest in the Age of “Post-Truth”. Conspiracy Theory in Turkey / J. de Medeiros. – Bloomsbury Publishing, 2018. – 222 p.; Douai A. Global and Arab Media in the Post-truth Era: Globalization, Authoritarianism and Fake News / A. Douai // IEMed: Mediterranean Yearbook. – 2019. – P. 124-132.

³⁶ Swartz L. Disability and Citizenship in the Global South in a Post-truth Era / L. Swartz // The Palgrave Handbook of Disability and Citizenship in the Global South / eds. B. Watermeyer, J. McKenzie, L. Swartz. – Cham: Springer International Publishing, 2019. – P. 57-65.

³⁷ Ilchenko S.N. Fake control, or News that should not be believed: How the media fool us / S.N. Ilchenko. - Phoenix, 2021. - 188 p. (In Russian).

³⁸ Romanova E.R. The persuasive power of fake news: the concept of cognitive distortion / E.R. Romanova // Journal of Dubna State University. Series: «Science of man and society». – 2021. – № 3. – P. 114-115. (In Russian).

³⁹ A systematic literature review on disinformation: Toward a unified taxonomical framework / E. Kapantai [et al.] // New Media & Society. – 2021. – Vol. 23. – № 5. – P. 1301-1326.

⁴⁰ Lewandowsky S. Beyond Misinformation: Understanding and Coping with the “Post-Truth” Era / S. Lewandowsky, U.K.H. Ecker, J. Cook // Journal of Applied Research in Memory and Cognition. – 2017. – Vol. 6. – № 4. – P. 353-369.

classifications exist on hate speech, as an alternative form of digital information threat, proposed by G. Pitruzzella⁴¹, E.A. Spiridovich⁴², and M.A. Paz and other authors⁴³.

As for the countermeasures, important contributions were made by A. Lohani⁴⁴, P. Fortuna⁴⁵, W. Timmerman⁴⁶ in the case of hate speech, R.K. Helm, H. Nasu⁴⁷, K. Dalkier⁴⁸, M.A. Alonso⁴⁹, and A.I. Cuza⁵⁰ in the case of fakes and disinformation. These studies highlight a wide range of initiatives covering both social and technical tools to combat false and malicious content.

The scientific understanding of information influence also includes other important studies, among which particular areas gained increasing importance. These include the study of the use of advanced technologies for manipulative purposes, as well as the general trend of “datafication” of international relations⁵¹. The central topic in this regard is artificial intelligence, or rather, its malicious use in world politics. E.N. Pashentsev⁵²

⁴¹ Pitruzzella G. *Disinformation and Hate Speech: A European Constitutional Perspective* / G. Pitruzzella, O. Pollicino. – Milano: EGEA Spa - Bocconi University Press, 2020. – 174 p.

⁴² Spiridovich E.A. Problems and the concept of hate speech in the Internet space / E.A. Spiridovich // *Electronic proceedings of young specialist works of the Polotsk State University. Legal Sciences.* - 2021. - Vol. 37. - P. 139-140.

⁴³ Paz M.A. Hate Speech: A Systematized Review / M.A. Paz, J. Montero-Díaz, A. Moreno-Delgado // *SAGE Open.* – 2020. – Vol. 10. – № 4. – P. 2; Internet, social media and online hate speech. Systematic review / S.A. Castaño-Pulgarín [et al.] // *Aggression and Violent Behavior.* – 2021. – Vol. 58.

⁴⁴ Lohani A. *Countering Disinformation and Hate Speech Online: Regulation and User Behavioural Change* / ORF [Electronic resource]. – URL: <https://www.orfonline.org/research/countering-disinformation-and-hate-speech-online/> (accessed: 20.12.2022).

⁴⁵ Fortuna P. A Survey on Automatic Detection of Hate Speech in Text / P. Fortuna, S. Nunes // *ACM Computing Surveys.* – 2018. – Vol. 51. – № 4. – P. 85:1-85:30.

⁴⁶ Timmermann W. Counteracting Hate Speech as a Way of Preventing Genocidal Violence / W. Timmermann // *Genocide Studies and Prevention: An International Journal.* – 2008. – Vol. 3. – № 3. – P. 353-374.

⁴⁷ Helm R.K. Regulatory Responses to ‘Fake News’ and Freedom of Expression: Normative and Empirical Evaluation / R.K. Helm, H. Nasu // *Human Rights Law Review.* – 2021. – Vol. 21. – № 2. – P. 302-328.

⁴⁸ Dalkir K. *Navigating Fake News, Alternative Facts, and Misinformation in a Post-Truth World* / K. Dalkir; ed. R. Katz. – 1st edition. – Hershey, PA: IGI Global, 2020. – 375 p.

⁴⁹ Sentiment Analysis for Fake News Detection / M.A. Alonso [et al.] // *Electronics.* – 2021. – Vol. 10. – № 11. – P. 1348.

⁵⁰ Farte G.I. Reactive Public Relations Strategies for Managing Fake News in the Online Environment / G.I. Farte, D.R. Obada // *Postmodern Openings.* – 2018. – Vol. 9. – № 2. – P. 26-44.

⁵¹ Tsvetkova N. A. Digital diplomacy and digital international relations: Challenges and new advantages / N. A. Tsvetkova, A. N. Sytnik, T.A. Grishanina // *Vestnik of Saint Petersburg University. International Relations.* – 2022. – Vol. 15. – № 2. – P. 174–196. (In Russian).

⁵² Pashentsev E. *Malicious Use of Artificial Intelligence* / E. Pashentsev // *European Conference on the Impact of Artificial Intelligence and Robotics. ECIAIR 2019.* – UK: Academic Conferences and Publishing International Limited, 2019.

and D.Y. Bazarkina⁵³, K.A. Pantserev⁵⁴, V. Bakir and A. McStay⁵⁵ are the key authors in this research area. The research on the malicious use of artificial intelligence is characterized by its multifaceted nature determining the mutual dependence of new information threats and the technical development of various actors.

In the context of theoretical and practical understanding of examples of countering information threats, resilience⁵⁶ and multistakeholderism⁵⁷ concepts occupy a special place, developing new aspects of the systemic response understanding to digital threats. In addition, numerous works analyze empirical data on the counter-disinformation practice at various levels of the international system. For example, the works of G. Pitruzzella⁵⁸, D.Y. Bazarkina⁵⁹, V.S. Carik⁶⁰, I. Nenadić⁶¹ reflect in detail the trends observed in the European Union, combining national practices of member-states with supranational regulatory intervention. At the same time, while these works have an extensive practical component, there is little (in some cases, a lack of) theoretical development or observation on how to conceptualize disinformation and respond to it.

⁵³ Pashentsev E. The malicious use of artificial intelligence against government and political institutions in the psychological area / E. Pashentsev, D. Bazarkina // *Regulating Artificial Intelligence in Industry*. – Routledge, 2021; Bazarkina D. Malicious Use of Artificial Intelligence / D. Bazarkina, E. Pashentsev // *Russia in Global Affairs*. – 2020. – Vol. 18. – № 4. – P. 154-177.

⁵⁴ Pantserev K.A. The Malicious Use of AI-Based Deepfake Technology as the New Threat to Psychological Security and Political Stability / K.A. Pantserev // *Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity : Advanced Sciences and Technologies for Security Applications* / eds. H. Jahankhani [et al.]. – Cham: Springer International Publishing, 2020. – P. 37-55.

⁵⁵ Bakir V. Op. cit. (Empathic Media...); Bakir V. Op. cit. (Fake News...).

⁵⁶ Tocci N. Resilience and the role of the European Union in the world / N. Tocci // *Contemporary Security Policy*. – 2020. – Vol. 41. – № 2. – P. 176-194; Resilience Concept in the European Union: Articulation and its Consequences for Russia / ed. T.A. Romanova. - St. Petersburg: Publishing House of the SPBU, 2019. (In Russian).

⁵⁷ Raymond M. Multistakeholderism: anatomy of an inchoate global institution / M. Raymond, L. DeNardis // *International Theory*. – 2015. – Vol. 7. – № 3. – P. 572-616; Antonova S. “Capacity-building” in global Internet governance: The long-term outcomes of “multistakeholderism” / S. Antonova // *Regulation & Governance*. – 2011. – Vol. 5. – № 4. – P. 425-445.

⁵⁸ Pitruzzella G. Op. cit.

⁵⁹ Bazarkina D. Evolution of the European Union’s Approach to Countering Disinformation / D. Bazarkina // *Vlast*. – 2021. – № 6. – P. 130-138. (In Russian); Bazarkina D. Evolution of Approaches to Countering Hybrid Threats in the European Union’s Strategic Planning / D. Bazarkina // *Contemporary Europe*. – 2021. – № 6. – P. 133-143. (In Russian).

⁶⁰ Tsarik V.S. Countering ‘Russian Information Threat’ In The European Union Policy After The Ukraine Crisis: Discourse And Institutional Aspects / V.S. Tsarik // *Central Russian Journal of Social Sciences*. – 2020. – Vol. 15. – № 5. (In Russian); Tsarik V.S. Fighting "Russian Disinformation" in the Public Positioning of Western Institutions: An Analysis of the Official Websites of NATO and the European Union / V.S. Tsarik // *Central Russian Journal of Social Sciences*. – 2019. – Vol. 14. – № 6. (In Russian).

⁶¹ Nenadić I. Unpacking the «European approach» to tackling challenges of disinformation and political manipulation / I. Nenadić // *Internet Policy Review*. – 2019. – Vol. 8. – № 4. – P. 1-22.

Some studies focus on particular institutions⁶² or areas of regulation, such as social media and platforms management⁶³, or countering the effects of the global COVID-19 pandemic on the information system⁶⁴. Important to mention is that research on information security and threat countering illustrates a general growth proportional to the activeness of various actors in risk mitigation associated with disinformation and hate speech.

The studies outlined above reflect two trends in information influence research. The first trend illustrates the diversification of threats and technical details behind modern digital information manipulation in international relations. The research subject develops a new terminological apparatus that needs further understanding. The second trend reflects the growing interest in various global, national, and regional counteraction practices. It leads to the creation of a separate research area related not to the content or the forms of information threats application, but to the social, regulatory, and political measures to tackle them. In this regard, this dissertation raises the **research question**: how a new terminological apparatus helps to understand and explain digital information threats in the 21st century, and what counteraction models to these threats emerge in world politics today? Based on the research question, the study makes the following **hypothesis**: the new terminological framework related to the “Post-truth” concept

⁶² Hedling E. Transforming practices of diplomacy: the European External Action Service and digital disinformation / E. Hedling // *International Affairs*. – 2021. – Vol. 97. – № 3. – P. 841-859; Vilson M. The Europeanization of Foreign Policy in the Face of the Russian Disinformation War / M. Vilson // *Sõjateadlane (Estonian Journal of Military Studies)*. – 2016. – № 2. – P. 114-140; Is the EU Disinformation Review Compliant with EU Law? Complaint to the European Ombudsman About the EU Anti-Fake News Initiative / A. Alemanno [et al.]. – Rochester, NY, 2018; Scheidt M. The European Union versus External Disinformation Campaigns in the Midst of Information Warfare: Ready for the Battle? *College of Europe EU Diplomacy Paper 01/2019* / M. Scheidt. – 2019. – 34 p.

⁶³ De Blasio E. Who Is Responsible for Disinformation? European Approaches to Social Platforms' Accountability in the Post-Truth Era / E. De Blasio, D. Selva // *American Behavioral Scientist*. – 2021. – Vol. 65. – № 6. – P. 825-846; Marsden C. Platform values and democratic elections: How can the law regulate digital disinformation? / C. Marsden, T. Meyer, I. Brown // *Computer Law & Security Review*. – 2020. – Vol. 36. – P. 105373; Saurwein F. Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe / F. Saurwein, C. Spencer-Smith // *Digital Journalism*. – 2020. – Vol. 8. – № 6. – P. 820-841; Kuczerawy A. Fighting Online Disinformation: Did the EU Code of Practice Forget about Freedom of Expression? / A. Kuczerawy // *Disinformation and Digital Media as a Challenge for Democracy*. – Intersentia, 2019. – P. 291-308; ; Wijermars M. Sociotechnical imaginaries of algorithmic governance in EU policy on online disinformation and FinTech / M. Wijermars, M. Makhortykh // *New Media & Society*. – 2022. – Vol. 24. – № 4. – P. 942-963.

⁶⁴ Current Response and Management Decisions of the European Union to the COVID-19 Outbreak: A Review / K. Goniewicz [et al.] // *Sustainability*. – 2020. – Vol. 12. – № 9. – P. 3838; Vériter S.L. Tackling COVID-19 Disinformation: Internal and External Challenges for the European Union / S.L. Vériter, C. Bjola, J.A. Koops // *The Hague Journal of Diplomacy*. – 2020. – Vol. 15. – № 4. – P. 569-582; Ivančík R. Fight Against Disinformation Concerning Coronavirus Pandemas At European Union Level / R. Ivančík // *Almanach (Actual Issues in World Economics and Politics)*. – 2021. – Vol. 16. – № 3. – P. 5-15.

complements, rather than complicates insight into new forms of information threats in world politics. It manifests in a parallel intensification of research and practical activities of international institutions in the development of multi-level and multifaceted countermeasures corresponding to the nature of the threat.

The scientific novelty of the research is built upon a combination of features related both to the research subject and to the applied theoretical and methodological framework. Firstly, the work establishes a systemic relationship between post-truth and the counteraction mechanisms for disinformation through the interrelation of the social environment and the system resources providing an expanded understanding of information influence characteristics in the digital age. Secondly, the author refers to the theoretical concepts of “Discourses of truth” and “Emotional regimes” allowing us to interpret information manipulations as a conflict of narratives, depending not only on the power struggle but also on the socio-cultural attitudes of society and the normative agenda of international institutions. Third, the study focuses on emerging threats by simulating potential and real examples of AI-based hate speech and disinformation. Fourthly, systemic modeling of information influence through the main actors’ behavior provides a new methodological tool for analyzing the actions of influence agents in world politics and the measures to counter them. Thus, this dissertation represents one of the first Russian studies of digital information threats theory in international relations concurrently relying on the “Post-truth” concept, the experience of global institutions, and a set of interrelated analytical and methodological approaches mentioned above.

The theoretical relevance of the study lies in the multi-level conceptualization of a new concept for the scientific community – “Post-truth”. The author seeks to exclude from the semantic field of post-truth those characteristics, that politicize and blur the concept, and to strengthen its applicability in the study of international relations. The theoretical analysis of the term is based on poststructuralist and postmodern ideas, which serve not as the cause of post-truth, but as its ideological forerunner. In addition, the study systematizes the main characteristics of post-truth politics as a modern condition for digital information influence. It consistently assesses the key gaps and targets for criticism inherent in this concept. Coherent responses to criticisms serve as the basis for

further theoretical application of “Post-truth” as a concept illustrating the “emotional” information environment of digital media.

The practical relevance consists of the possible application of the proposed theoretical and methodological developments, as well as the developed countermeasures by public authorities of the Russian Federation in the development of the theoretical sections of concepts and doctrines in the fields of information security or the countering of disinformation and hate speech proliferated by unfriendly countries and other actors seeking to harm Russia. Numerous provisions are applicable in the educational process on “International relations” and for conducting fundamental scientific research.

The structure of the dissertation corresponds to the described logic of the study and includes an introduction, three chapters, concluding remarks, a list of sources and bibliography, and appendices. Each chapter is divided into subchapters corresponding to the research tasks. The first chapter describes the initial philosophical and theoretical basis of the work, considering the categories of “Truth” and “Falsehood” through the lens of poststructuralism and postmodernism; reflects a bibliographic analysis of the “Post-truth” concept and evaluates the scientific criticism of the concept. The second chapter defines the methodological framework of the work. It structures the main aspects of the information influence analysis model proposed by the author; consistently presents the characteristics of the subject, object, and intermediary; establishes the tools for studying each of the actors. The second chapter provides key information threats, disinformation, and hate speech classifications. The third chapter analyzes future scenarios for improving threats through artificial intelligence. It also presents the main characteristics of current experience in countering information threats by actors and institutions at various levels of the international system. The chapter assesses the resilience and multistakeholderism concepts as promising ideas in information regulation. The conclusion summarizes the study's results and assesses the prospects for further digital information influence studies. The appendices reproduce the visual supplement to the information influence studying model proposed by the author of the dissertation.

The conclusions and ideas formulated in the work were **tested** in a number of publications, as well as the author's presentations at scientific events. The total number of

publications on the topic: 15 (including 7 scientific articles, 2 chapters in collective monographs, and 6 conference proceedings – see Sources and Bibliography: Author’s papers prepared in the course of writing the dissertation). 5 publications are indexed in journals recommended by the Higher Attestation Commission, 4 are indexed in the international Web of Science and Scopus databases. Separate parts of the study were financially supported by research grants from the Russian Science Foundation (RSF)⁶⁵ and the Russian Foundation for Basic Research (RFBR)⁶⁶.

The **main scientific results** achieved during the dissertation research include:

1) A theoretical understanding of digital threats (disinformation and hate speech) through the prism of the concept of “Post-truth,” taking into account various social⁶⁷, discursive⁶⁸, and macro-political⁶⁹ variables. In particular, the research justified the concept’s scientific usefulness to contemporary international relations research⁷⁰.

2) The creation of a conceptual classification model of measures to counter information threats through the systemic response logic of various actors of the international system to external and internal triggers, among which the paternalistic and adaptive response to information threats⁷¹ play a central role.

⁶⁵ The author of the dissertation participated in the research grant “Resilience concept of the European Union: articulation and its implications for Russia”. RSF (Project No. 17-18-01110). March - December 2019

⁶⁶ The author of the dissertation participated in the research grant “Malicious Use of Artificial Intelligence and Challenges to Psychological Security in Northeast Asia,” RFBR (Project No. 21-514-92001). 2021 - 2022

⁶⁷ Kolotaev Y. Political Implications of Hate Speech Digitalization in a Post-Truth Era: Impact on Emotional Regimes in Digital Conflicts / Y. Kolotaev // RUDN Journal of Political Science. - 2022. - Vol. 24. - №3. - P. 520. (In Russian); Kolotaev Y. Post-truth through the lens of emotional regimes and emotives // Media in the modern world. 61st Petersburg Readings. St. Petersburg, 2022 - P. 106. (In Russian)

⁶⁸ Kolotaev Y.Y. Russia’s identity formation through the process of its interactions with the European Union within the context of post-truth politics / Y. Kolotaev // Russia in the Global World: New Challenges and Opportunities. – 2018. – P. 103. (In Russian).

⁶⁹ Kolotaev Y.Y. Fake news, disinformation and strategic communications as key categories in EU-Russia relations within a post-truth context. External threats and internal challenges / Y. Kolotaev // Russia in the Global World: New Challenges and Opportunities. – 2019. – P. 432. (In Russian).

⁷⁰ Kolotaev Y. Prospects and limitations of the scientific use of the post-truth concept / Y. Kolotaev // Future World, Common Efforts, Real Progress. Proceedings of the International scientific-practical conference of young scientists. – Moscow, 2020. – P. 86. (In Russian).

⁷¹ *The model has been presented in the following publications:* Op. cit. (Resilience Concept in the European Union ...). P.93-95. *Personal contribution: Chapter § 3.2. Fake news or disinformation;* Kolotaev Y.Y. European Union in the age of post-truth: Developing societal resilience before European Parliament elections 2019 / Y.Y. Kolotaev // Vestnik of Saint Petersburg University. International Relations. – 2020. – Vol. 13, Issue 1. – P. 79; Romanova T.A. Disinformation (fake news, propaganda) as a threat to resilience: approaches used in the EU and its member state Lithuania / T. A. Romanova, N. I. Sokolov, Y.Y. Kolotaev // Baltic Region. – 2020. – Vol. 12. № 1. – P. 55-56. *Personal contribution – 15/5, Paradigms of countering disinformation (fake news) and resilience; Three approaches to disinformation in a comparative perspective.*

3) The elaboration of the new media actorness problematics in the system of digitalizing international relations in the context of the growing influence possibilities of online platforms on the states' political environment⁷². Among the main observations in digital agents' activity, the research shows a trend toward multistakeholderism as a method of non-state actors' involvement in the digital sphere governance problems⁷³.

4) The actualization of the scientific understanding of prospective digital threats in information influence with the support of AI systems⁷⁴. The results presented a model for analyzing the key vulnerabilities of social and technical nature in information manipulation in world politics⁷⁵ and the structure of an AI-enhanced information attack⁷⁶.

5) A comprehensive analysis of the existing international experience in combating disinformation⁷⁷ for further theoretical structuring of the collected practice to form the basis for new methods of information regulation.

The conclusions to be defended include the following observations and findings from the study:

- In modern conditions, a relevant and heuristically useful interpretation of the devaluation of truth in the context of informatization and the spread of social media is the concept of "Post-truth politics". It combines technical premises with the problem of social polarization in the realities of the 21st century. Post-truth emphasizes the practical importance of emotions while determining the veracity of statements and judgments, or while exerting information influence. "Post-truth" has structural problems associated with the political and journalistic aspects of its articulation, which need to be eliminated. The

⁷² Kolotaev Y., Kollnig K. Political influence of online platforms: YouTube's place in European politics / Y. Kolotaev, K. Kollnig // Vestnik of Saint Petersburg University. International Relations. – 2021. – Vol. 14. – №. 2. – P. 228. *Personal contribution: 10/5, Power and actorness of mass and social media; Types of political influences.*

⁷³ Kolotaev Yu. Yu. Evolution Of The EU's Policy Of Tackling Disinformation. A Multistakeholder Approach / Yu. Yu. Kolotaev // RSUH/RGGU Bulletin Series "Political Science. History. International Relations." – 2023. – № 2. – P. 50. (In Russian)

⁷⁴ Kolotaev Y. Methods for Countering the Malicious Use of Artificial Intelligence: Northeast Asian Experience and Models / Y. Kolotaev // Information wars. – 2022. – №. 3. – P. 65. (In Russian)

⁷⁵ Kolotaev Y. Sentiment Analysis: Challenges to Psychological Security and Political Stability / Y. Kolotaev // Proceedings of the 3rd European Conference on the Impact of Artificial Intelligence and Robotics ECIAIR 2021, 2021. P. 85.

⁷⁶ Kolotaev Y. Hate Speech in Perception Management Campaigns: New Opportunities of Sentiment Analysis and Affective Computing / Y. Kolotaev // The Palgrave Handbook of Malicious Use of AI and Psychological Security / ed. E. Pashentsev. – Cham: Palgrave Macmillan, 2023. – P. 116.

⁷⁷ Kolotaev Y. Malicious Information Threats of a Post-Covid World in Europe and the European Union // Contemporary Europe. – 2021. – № 7. – P. 145–146. (In Russian); Kolotaev Y. Disinformation in Times of Pandemic: Experience of the EU in Tackling New Digital Threats / Y. Kolotaev // The world after the pandemic: global challenges and prospects for development. Global Development, Iss. 23. IMEMO, 2022 – P. 114-115. (In Russian).

resources for addressing these problems lie within the interdisciplinary nature of the concept, the emphasis on technical transformations, and particular attention to the construction of truth through emotions as social interaction. Further elaboration of these aspects will allow “Post-truth” to secure the status of an independent theoretical and philosophical category applicable to interstate and international relations.

- The main elements that form the structural basis of information influence are the links between the subject, object, and intermediary. The central connecting element is the interrelation of the subject and the object in their inherent behavioral logic. For the subject, a rational choice model determines the form and content of a chosen strategy of behavior. The internal logic of the object of influence depends on the level of influence: social or individual. At the individual level, rational thinking meets bounded rationality, reflecting a system of fast cognitive mechanisms allowing one to respond rapidly to incoming information. At the social level, truth discourses and emotional regimes are responsible for the reaction system, generating collective consensus on the veracity of individual facts and the acceptability of specific social emotions.

- The digitalization of the media (intermediary) leads to the simplification of information influence through new emerging vulnerabilities and mechanisms of the intermediary. Post-truth accelerates the spread of unverified facts through the emotionality of the discourse formed by political actors. For this purpose, social network algorithms are misused and exploited. However, the political context is not the only condition for creating disinformation and hate speech. They also emerge under the influence of situational, non-political, and subjective factors. This only complicates the creation and implementation of necessary measures to counter digital threats.

- The use of artificial intelligence in digital manipulations increases currently potential and real threats of information influence. The technical optimization of disinformation includes the exploitation of algorithms for content creation while spreading fakes, the creation of synthetic data and information, and the use of digital emotional analysis of the influence object for empathic optimization of information.

- The existing practices of countering information threats can be categorized according to the levels of the international system, i.e., global, supranational, national,

regional, and local measures. Overall, these practices reflect one of two different governance logics. One includes adaptation to systemic pressure, using new challenges as an opportunity for the system's evolution. The other relies on a paternalistic vision, i.e., protective measures aimed at maintaining the system's existing balance or rebuilding it. Various combinations of these two governance logics indicate a large variability of practices, which in sum form a multistakeholder approach build on the cooperation of various actors. It incorporates horizontal and vertical cooperation to develop best practices.

Chapter 1: The notions of “Truth” and “Post-truth” in international security: A perspective on digital information influence and online manipulations

The current political situation at the global and local levels is associated with a growing conflict in international relations, information wars, the evolution of populism, and a decline in trust in expert knowledge. The transformation of the information space, taking place in parallel with these processes, dealt a significant blow to the public consensus regarding the existing system of production, consumption, and evaluation of information. In the context of the growing variety of information flows capable of carrying any information and unlimited access to them, humanity faces a new, modified interpretation of the previously existing phenomena in the communicative space. Lies and disinformation, as natural features of social and interpersonal communication, have acquired a new meaning in terms of “Fake news” and “Alternative facts”, and structured information influence, similar to propaganda, has come to be called “Strategic communications”. The change of concepts and the growth of their use characterized the changes in the perception of the conditions in which society finds itself. The central place belongs to the decrease in the role of facts and an increase in the influence of emotions. A new idea about the spread of “Post-truth politics” or the “Post-truth era”, used to denote a change in society’s attitude towards the factual content of information and its rational perception, has penetrated the political discourse of various countries.

In the social sciences, the spread of new information threat types (fakes, hate speech, etc.) caused by informatization is connected precisely with the circulation of post-truth conditions. However, despite the novelty of information threats, a significant part of them has a set of factors rooted in cognitive psychology relevant throughout human history. So, for example, the phenomenon of disinformation, even in the modern technological embodiment, is closely connected with the understanding of such historically universal entities as “Truth”, “Reliability”, and “Objectivity”. Each concept is a fundamental aspect of philosophical thought at all stages of its formation and development. However, all these concepts (“Truth”, “Reliability”, and “Objectivity”)

received new interpretations in the 20th and 21st centuries, in particular, within the framework of postpositivist philosophical thought.

For this reason, the philosophical and theoretical study of new information threat types is highly significant, allowing us to connect the ongoing changes with the ideas already at the disposal of scientists and thinkers. In the context of fake information and post-truth, the concepts of “Truth” and “Falsehood” receive a new reading and interpretation. At the same time, of particular interest here is the understanding of these categories by representatives of postmodernism and postpositivism, who have changed the modern worldview in the field of the problem of (in)comprehensibility of truth and the reality of the world around us, as well as in understanding information security through the prism of discourse and narrative.

Before starting the chapter, it is important to remark on the philosophical and political analysis of the concept of “правда” (“truth” in Russian) and the problem of its boundaries in the Russian language in its relationship with the very close meaning term “истина” (semantically close, but not interchangeable to “verity”). Despite the notion of the semantic boundaries of each of the concepts that exist in Russian philosophy⁷⁸, within the framework of this work, both terms are generalized to avoid conceptual confusion in connection with the translation in Russian of the English term “post-truth” (“постправда” or “постистина”⁷⁹ in Russian), whereby the well-established translation in Russian is “постправда”. This need is due to intersections in the interpretation of the word “truth” in separate concepts that have different understandings in the Russian language in their relation to “истина” and “правда” (e.g, the use of M. Foucault’s “discourses of truth” for the study of post-truth)⁸⁰. The generalization and partial interchangeability of both terms in this study aim to overcome false interpretations and linguistic divergence since, in the

⁷⁸ Karasika V.I., Sternina I.A. Anthology of concepts / Ed. V.I. Karasika, I.A. Sternin. Volume 3. Volgograd: Paradigm, 2006. - 381 p. (In Russian); Chernikov M.V. The Truth and the Varsity Concepts in Russian Culture: the Problem of Correlation / M.V. Chernikov // Polis. Political Studies. – 1999. – № 5. – P. 43-61. (In Russian); Lishaev S.A. "Правда" and "истина" (linguistic conceptualization of the world and thematic originality of Russian philosophy) / S.A. Lishaev // Bulletin of the Samara Humanitarian Academy. – 2006. – № 1 (4). – P. 173-209. (In Russian).

⁷⁹ Morozov A. How “Post-Truth World” Finally Became Hyperstition / A. Morozov // Philosophy. Journal of the Higher School of Economics. – 2021. – Vol. 5. – № 3. – P. 287-297. (In Russian); Salin A. Post-truth and Its Threats / A. Salin // Russian sociological review. – 2022. – Vol. 21. – № 1. – P. 284-297 (In Russian).

⁸⁰ Foucault M. Subjectivity and Truth: Lectures at the Collège de France, 1980-1981 : Michel Foucault, Lectures at the Collège de France. Subjectivity and Truth / M. Foucault; eds. F. Ewald, A. Fontana, F. Gros. – London: Palgrave Macmillan UK, 2017. – 331 p.

political research on “post-truth”, the linguistic differences of truth give way to the opposition of truth to lies and facts to emotions.

1.1 Truth, power and information influence through the lens of postpositivist and postmodernist epistemology

From a historical perspective, information manipulation exists as a means of political and military (as well as economic, etc.) influence at the international, state, and public levels no less than the system of social organization of people itself. The phenomenon of information influence is associated with a whole range of concepts and terms that sometimes have different meanings but are united by a common thematic direction. This complex includes established (disinformation, propaganda) and more modern notions (fake, factoid). The study of the semantic and theoretical properties of each of these concepts is the subject of multiple scientific works⁸¹ that seek to identify or refute the novelty of the threats to information security experienced in modern times associated with *the distortion or deliberate interpretation of information in a politically beneficial way for any social group*.

At the same time, technological development in the 21st century becomes the main incentive for a new iteration of understanding distortions and lies in the information space. The emergence of *new means of disseminating information* transforms the perception process, and hence the forms of manipulation (close in meaning are M. McLuhan’s ideas that “the medium is the message”)⁸². At this junction of the roots of disinformation that go deep into human history and the novelty of this phenomenon that constantly arises from technological progress, there is the problem of “*Truth*” as a philosophical category. Regardless of the historical conditions and the state of social relations, it is the truth falling under the direct influence of “Fakes”, meaning reflections of reality, which become distorted, changed, or turned into a lie as a result of information influence.

⁸¹ Tandoc E.C. Defining “Fake News”: A typology of scholarly definitions / E.C. Tandoc, Z.W. Lim, R. Ling // Digital Journalism. – 2018. – Vol. 6. – № 2. – P. 137-153.; Farte G.I. Op. cit.

⁸² McLuhan M. Op. cit.

At its core, the distortion of the truth is a phenomenon difficult to interpret, especially because the main problem is the comprehension of the very category of “Truth”. The development of philosophical thought in the 20th century changed the classical idea of the Enlightenment about truth and fact as the unshakable basis of human knowledge of the world. A special place within the framework of such philosophical thought belongs to postpositivism, which is not so much a single coherent theory as a collective term for those theoretical views skeptical of positivist views on knowledge, science, facts, and the possibilities of ultimate knowledge of reality⁸³. For this reason, postpositivism in the broadest context includes postmodernism, critical theory, poststructuralism, and other approaches that express doubts to varying degrees about the universality of the category “Truth” and have spread beyond philosophy to other social sciences, including the study of international relations⁸⁴.

Postpositivism brought into the theory of international relations the belief that any event and phenomenon is an object of *interpretation*, depending on the specific situation, context, linguistic means, and moral attitudes. From a security perspective, D. Campbell claims that “nothing is intrinsically more dangerous [...] than anything else, except when interpreted as such.”⁸⁵ The presence of such a political dimension in the act of interpretation makes determining “truthfulness” in politics a security issue or, more broadly, an instrument of power.

Under the influence of the ideas of postpositivist philosophy, the Western academic community began a critical rethinking of such fundamental concepts as “Truth”, “Rationality”, “Objectivity”, “Reality”, and “Consensus”⁸⁶. For postpositivists, what has become most important is the study of “Truth” by identifying the social groups and interests behind *truth claims*, as well as how such claims can influence truth⁸⁷, which

⁸³ Making sense of international relations theory / ed. J.A. Sterling-Folker. – Second edition. – Boulder, Colo: Lynne Rienner Publishers, Inc, 2013. – 481 p.

⁸⁴ Ibid.

⁸⁵ Campbell D. Writing Security: United States Foreign Policy and the Politics of Identity / D. Campbell. – Revised edition. – Minneapolis: University of Minnesota Press, 1998. – P.2.

⁸⁶ Fay B. Theory and metatheory in social science — or, why the philosophy of social science is so hard / B. Fay // *Metaphilosophy*. – 1985. – Vol. 16. – № 2/3. – P. 150-165; Lapid Y. The Third Debate: On the Prospects of International Theory in a Post-Positivist Era / Y. Lapid // *International Studies Quarterly*. – 1989. – Vol. 33. – № 3. – P. 235-254.

⁸⁷ Fluck M. The concept of truth in international relations theory: critical thought beyond post-positivism. The concept of truth in international relations theory / M. Fluck. – London, United Kingdom: Palgrave Macmillan, 2017. – 247 p.

means that it is impossible to separate the process of studying politics from the norms, practices, and institutions linked to politics. As a result, the subject and object of research merge, making the creation and interpretation of truth a matter of power relations.

A similar trend within the social sciences and political science indicates that there is a “relativization of philosophical thinking has greatly complicated the task of providing effective legitimation of knowledge”⁸⁸. But, despite such relativism and perspectivism⁸⁹, the postpositivist approach introduced a new understanding of truth, which represented the clash of the points of view of poststructuralism and critical theory (both theoretical approaches broadly adhere to postpositivist epistemology). So, in postpositivism, a negative and positive understanding of the truth has developed⁹⁰. Inherent in the poststructuralist interpretation of truth (Campbell’s “National Deconstruction” is an example here) is an emphasis on the political significance of truth as a source of domination and violence⁹¹. An alternative to the negative perception of truth is the interpretation of critical theory, which emphasizes the intersubjective understanding of truth. A. Linklater, in his work “The Transformation of Political Community”, emphasizes the importance of discursive consensus, which makes it possible to make an intersubjective understanding of truth a source of progress⁹².

At the same time, the negative and positive understanding of the truth has two general provisions. The first represents a single perception of truth as a “social phenomenon”, and the second sees the truth through “intersubjective epistemological practices”, emphasizing the *conditions* for creating truth⁹³. Both of these views make it prove that any attempt to identify and consolidate the truth (whether negatively or positively) is a collective action that depends to a large extent on external and internal political circumstances.

⁸⁸ Lapid Y. Op. cit. P. 243.

⁸⁹ Halliday F. Book review: a “crisis” of international relations? / F. Halliday // *International Relations*. – 1985. – Vol. 8. – Book Reviews. – № 4. – P. 412.

⁹⁰ Fluck M. Op. cit. P.13.

⁹¹ Campbell D. *National deconstruction: Violence, identity, and justice in Bosnia* / D. Campbell. – Univ of Minnesota Press, 1998. – 328 p.

⁹² Linklater A. *The transformation of political community: ethical foundations of the post-Westphalian era. The transformation of political community* / A. Linklater. – Reprinted. – Cambridge: Polity Press, 2004. – 263 p.

⁹³ Fluck M. Op. cit. P.13.

This understanding of the truth leads to an acute (for the international context) possibility of (external and internal) information influence on the truth creation process. From the point of view of intersubjectivity, any relevant discursive practice (even misinformation) can become an integral part of setting the truth. In this case, any falsity might embed in a factual statement while being created. Falsity itself can be perceived as an interpretation of reality, depending on the institutions involved in truth creation, thus, becoming an integral part of the truth recognized by any part of society. This opportunity is exploited in the deliberate dissemination of disinformation and embodies the term “Factoid”⁹⁴. In this matter, the (un)truth becomes, as reflected in its negative understanding, a source to dominate because politically motivated information influence leads to the social recognition of lies as truth in the absence of any significant opposition to this from other political actors.

Such a postpositivist understanding defines truth not in terms of reflecting the “real world” but as a social consensus. The agreement, resulting from the intersubjective acceptance of lies at the national or international level as the truth, allows us to explain the reasons for the possible achievements of information influence in public discourse. In this case, important is not what is true but what becomes a social practice. Consequently, the boundaries between lies and truth are blurred and reduced to questions of political influence and the ability of political actors to reinforce or, conversely, expose any statement that claims to be true.

On the one hand, this logic allows us to point out the danger of any information influencing attempts on the public (because they can transform the consensus set in society about some phenomenon or fact). On the other hand, the question arises connected to the postpositivist epistemology about the existence of truth per se outside of intersubjective interpretations. In the context of relativistic epistemology, this issue was, in particular, considered by one of the most influential poststructuralist philosophers, M. Foucault. In his lectures on subjectivity and truth, Foucault refers to the “Discourses of

⁹⁴ Definition of FACTOID [Electronic resource]. – URL: <https://www.merriam-webster.com/dictionary/factoid> (accessed: 06.02.2023).

truth”⁹⁵ (in Russian also “дискурсы истины”)⁹⁶. In his understanding, the existence of a discourse that claims to denote the truth does not mean that there is a direct connection between the discourse and the reality that he speaks of⁹⁷. Foucault questions why truth exists in addition to reality⁹⁸. For him, the world’s reality is not a truth as such. Discourses of truth do not only document what is part of reality but are themselves a fragment of reality that does not need to appear separately⁹⁹. Truth, therefore, is a circumstance not obligatory for the reality, but supplementing it in case of its appearance.

Within discourses of truth, the truth becomes an element of changing reality and can not fully reflect it. As with knowledge systems influenced by power relations, truth is the product of historical circumstances¹⁰⁰. Therefore, it cannot refer to the fundamental notion that there is a single truth. A plurality of truth becomes an alternative perspective.

Thus, Foucault, referring to the discourses of truth, does not deny the existence of truth itself. For him, its origin, which is often random and optional¹⁰¹, is of greater importance. At the same time, despite its optionality, truth is a crucial aspect of governmentality¹⁰² for Foucault. The discourse of truth underlies governance because of the limited access to the possibility of proclaiming the truth. It cannot be distributed equally in society since “not everyone can tell the truth just because everybody may speak.”¹⁰³ Access to the process of forming the truth has a social framework, which turns it into a governance tool while maintaining its dynamic and polymorphic nature.

Foucault’s point of view, associated with limited access to the creation of truth, partly contradicts the idea of truth as an intersubjective social phenomenon. However, such a position makes it possible to narrow and identify the circle of potential actors

⁹⁵ Foucault M. Op. cit. (Subjectivity and Truth...).

⁹⁶ Grozina N.A. The problem of the source of power and the theme of resistance in Foucault's philosophy / N.A. Grozina // Vestnik NSU. Series: Philosophy. - 2006. - Vol. 4. - No. 2. - P. 153-159. (In Russian); Hair J. Religion and Morality (In Russian) / Stanford Philosophical Encyclopedia: Translations of Selected Articles [Electronic resource]. – URL: <https://philosophy.ru/ru/religion-morality/index.php> (date of access: 02/24/2023).

⁹⁷ Foucault M. Op. cit. (Subjectivity and Truth...). P. 221.

⁹⁸ Ibid. P. 237

⁹⁹ Prozorov S. Why is there truth? Foucault in the age of post-truth politics / S. Prozorov // Constellations. – 2019. – Vol. 26. – № 1. – P. 18-30.

¹⁰⁰ Foucault M. Nietzsche, Genealogy, History / M. Foucault // The Foucault reader / ed. P. Rabinow. – New York: Pantheon Books, 1984. – P. 76-100.

¹⁰¹ Prozorov S. Op. cit.

¹⁰² Foucault M. The Government of Self and Others: Lectures at the Collège de France 1982–1983. The Government of Self and Others / M. Foucault; eds. F. Gros, F. Ewald, A. Fontana. – London: Palgrave Macmillan UK, 2010. – 402 p.

¹⁰³ Ibid. P. 183.

having the real opportunity to transform and determine the truth. In the case of information influence and disinformation, the connection of truth with access to control means that the ability to influence discourse is a matter of politics and power.

Despite the initial narrowness of actors of this process, changes in the digital environment signal a reverse trend towards decentralization of the discourses of truth, i.e., the meaning of individual discursive practices becomes blurred in the information flow. Therefore, along with the change in the conditions for truth dissemination, a revision of the main stakeholders of truth(s) creation.

The application of discourses of truth in the formation of reality (the instrumentalization of truth) is the main prerequisite for disinformation and other information threats. However, to understand the current situation in this area, besides considering the essence of truth as a social phenomenon, it is necessary to turn to a theoretical understanding of the philosophical characteristics of modernity. Within the framework of postpositivist epistemology, the ideas reflected in the postmodern philosophy are especially relevant.

Being an integral part of a broader movement, postmodernism, like all related postpositivist epistemology¹⁰⁴, is very diverse and difficult to categorize. The problem in assessing postmodernism derives from its desire to overthrow any preceding epistemology¹⁰⁵. In contrast to critical theory, which mainly considers exclusion and dominance issues, postmodernists are united by doubt about the possibility of producing any objective and rational knowledge¹⁰⁶.

Within the framework of postmodernism, there are two main ideas in the perception of the events of the modern world. The first indicates that any event or challenge in world politics does not precede our views and theories about it ontologically¹⁰⁷. Basic ideas and prevailing ways of thinking and acting predetermine reality and become reproduced in it.

¹⁰⁴ Op. cit. (Making sense of international relations theory); Bleiker R. Postmodernism / R. Bleiker // Introduction to international relations: Australian perspectives / eds. R. Devetak, A. Burke, J. George. – Cambridge ; New York: Cambridge University Press, 2007. – P. 86-95.

¹⁰⁵ International theory: positivism and beyond / eds. S. Smith, K. Booth, M. Zalewski. – Cambridge ; New York: Cambridge University Press, 1996. – 362 p.

¹⁰⁶ Wight C. Philosophy of Social Science and International Relations / C. Wight // Handbook of International Relations. – London: SAGE Publications, 2002. – P. 23-51.

¹⁰⁷ Op. cit. (International theory: positivism and beyond).

The second characteristic is the focus on the decision process, what is an essential issue of world politics, and what is not¹⁰⁸ (a problem close to the essence of exclusion and domination, raised in the critical theory). These two circumstances pose questions to postmodern philosophy on our understanding and knowledge of political reality and its interpretation and reproduction. The circle of authors considering these issues includes the most famous representatives of postmodernism, for example, Lyotard and Baudrillard, along with philosophers having a significant impact on it (Foucault, Derrida, Rorty).

Assessing the state of reality, J.-F. Lyotard designated the existing situation of postmodernity as circumstances in which, influenced by progress, a distrust of metanarratives¹⁰⁹ appears. With the disuse of metanarratives, “the narrative function loses its functors [...]. It disperses into clouds of linguistic [...] particles, each carrying a pragmatic sui generis valency. Each of us lives at the intersections of the trajectories of many of these particles.”¹¹⁰ To clarify the features of such a situation, Lyotard refers to the concept of “Language games” (developed by Wittgenstein)¹¹¹ and gives it a new interpretation in postmodern conditions. Such language games have a “local determination” and a multifaceted character.

Even with a clear emphasis on the linguistic component, Lyotard points to some of the political aspects of postmodernity, which make it possible to introduce an element of technological determinism, revealing the influence of technology on the perception of reality. Describing the process of informatization of society in the context of the development of modern science, Lyotard notes that informatization “can become a 'desirable' tool for controlling and regulating the system on the go, extending to the control of knowledge itself.”¹¹² Amid informatization, the status of knowledge is also changing, fixing itself in the structure of competition for power in the international arena. Lyotard reflects this thought in the statement that “in the form of an information

¹⁰⁸ Op. cit. (International theory: positivism and beyond).

¹⁰⁹ Lyotard J.-F. *The Postmodern Condition* / J.-F. Lyotard; transl. N. A. Shmatko. - St. Petersburg: Aleteyya, 1998. - 160 p. (In Russian).

¹¹⁰ Ibid. P. 10-11.

¹¹¹ Wittgenstein L. *Philosophical investigations* / L. Wittgenstein. - Oxford: Basil Blackwell, 1968. - 250 p.

¹¹² Lyotard J.-F. Op. cit.

commodity necessary for increasing productive power, knowledge is already and will be the most important, and perhaps the most significant stake in the world rivalry for power.”¹¹³

Thus, the loss of metanarratives in the context of informatization and global information rivalry is a situation in which it is possible to exploit for information influence *the plurality of local narratives, the power measurement of information and knowledge*. The emerging pluralism also implies the replacement of the criterion of truthfulness (both in science and society) with performativity¹¹⁴. Effectiveness and practical significance of knowledge affect the information received, forcing, in the conditions of the fall of metanarratives, to turn knowledge into an “information product”. This point of view on information and its political aspects characterizes the postmodern.

Changes in the characteristics of information and its relationship with reality play a big role in the work of another significant postmodernist, J. Baudrillard. Being a popularizer of two crucial concepts, “Hyperreality” and “Simulacra”, he expressed in his work “Simulacra and Simulations” the idea that the modern world consists of a set of models that do not have any origins in reality and do not refer to it, but only to themselves. The *simulacra* are “the truth that hides that it does not exist.”¹¹⁵ They establish through simulation a model of reality without an original, or, in other words, a hyperreality¹¹⁶. Such a reality is self-replicating through simulations, becoming central to explaining modernity.

Simulation, according to Baudrillard, “differs from fiction or falsehood in that it not only presents the absence as presence, and the imaginary as real but also in that it undermines any contrast to reality by absorbing the real into itself.”¹¹⁷ Simulation logic puts the interpretation of the event before the event itself (*precession of simulacra and models*). The explanation of an event, the search for evidence, and its connection to reality, in the absence of any reference point, lead to the equivalence of any

¹¹³ Lyotard J.-F. Op. cit. P. 20.

¹¹⁴ Anderson P. The origins of postmodernity / P. Anderson. – London ; New York: Verso, 1998. – 143 p.

¹¹⁵ Baudrillard J. Simulacra and simulation / J. Baudrillard. - Moscow: Postum, 2017. - 240 p. (In Russian).

¹¹⁶ Ibid.

¹¹⁷ Baudrillard J. Selected writings / J. Baudrillard, M. Poster. – Stanford, Calif: Stanford Univ. Press, 2007. – 294 p.

interpretations. An important phenomenon in this context is “the *political* problem of imitation, hypersimulation, and aggressive simulation.”¹¹⁸ Without a clear division of reality and illusion, there is a need for a political response to phenomena.

Violation of the boundaries between reality and illusion delivers a direct blow to power and order. Power, relying on the political governance of the real, is in crisis. Under such conditions, Baudrillard states the illusion of power forming its own simulation: “In fact, power exists today only to hide that it no longer exists.”¹¹⁹ Thus, the lack of power becomes the main prerequisite for information uncertainty.

Together with the fall of metanarratives outlined by Lyotard, the idea of hyperreality demonstrates that reality is in an absolute pluralism of interpretations. They are an integral part of political reality since interpretation, simulations, and truth are all part of a single discursive space, unable to distinguish between fiction and truth because of the absence of actual (rather than simulated) power within such a discourse. Such conditions become a fertile environment for the spread of internal disinformation in a state or information interventions, the authenticity of which cannot be proved or disproved. *The desire to counter disinformation becomes a simulacrum per se* because it refers to a phenomenon whose reality cannot be proven.

The central role in the blurring of truth and reality belongs to the sources of information dissemination in the postmodern era. “All the media and official sources of information are applied only to maintain the illusion of events, the reality of goals, the objectivity of facts.”¹²⁰ An equally important remark by Baudrillard is that “the media carry meaning and counter-meaning, they manipulate in all directions at once, no one can control this process.”¹²¹ If we turn to the experience of the 21st century, most of Baudrillard’s statements correspond to the structure of modern digital media (especially social networks). At the same time, the simulation effect strengthened due to the mass character and virality of new information in the changed communication environment.

¹¹⁸ Baudrillard J. Op. cit. (Simulacra and simulation).

¹¹⁹ Ibid. P. 40.

¹²⁰ Baudrillard J. Op. cit. (Simulacra and simulation). P. 59.

¹²¹ Ibid. P. 117.

Baudrillard actively criticizes information as a form of communication, noting that it destroys its own content and social property because, with an increase in its amount, there is a proportional decrease in its meaning. “Instead of being the height of communication, information exhausts its power in staging communication” and meaning¹²². Imaginary interactivity and staging of communication intensification become interrelated phenomena. Staged communication decomposes the social, leading to its fragmentation and the loss of the function of an intermediary by media.

In modern discourse, situations of this kind have been called “Filter bubbles”¹²³ and “Echo chambers.”¹²⁴ These circumstances of distorted reality create, by themselves, a space in a state of informational impact on reality, adjusting it to the audience. These discrete communities within the dysfunctional media enable various political actors to use the most targeted disinformation since, in a staging communication, the line between disinformation and objective information blurs for the audience.

The violation of the communicative function of information also manifests itself in its break with the temporal aspect. It is indicated in the later works of Baudrillard in the thesis: “The closer we are to real-time and live broadcast, the further we will move in this direction.”¹²⁵ The “real-time utopia” is driven by the illusion of simultaneity, which requires constant interpretation and generates excessive comments with a lack of images¹²⁶. As a result, events and “non-events” have equal value because they produce interest and interpretation equally, and, thus, “the consequences of what did not happen can be as significant as the consequences of a real historical event.”¹²⁷ The political reality of the 21st century demonstrates this idea especially clearly. Unverified and unproven evidence has more than once become the stimulus for a reaction expressed in direct political action at the local or global levels. It relates to conflict situations in various

¹²² Baudrillard J. Op. cit. (Simulacra and simulation).

¹²³ Pariser E. The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think. The Filter Bubble / E. Pariser. – Penguin, 2011. – 179 p.

¹²⁴ Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? / P. Barberá [et al.] // Psychological Science. – 2015. – Vol. 26. – № 10. – P. 1531-1542.

¹²⁵ Baudrillard J. The Spirit of Terrorism. There was no Gulf War / J. Baudrillard; transl. A. Kachalov. - Moscow: RIPOL classic, 2016. - P. 47. (In Russian).

¹²⁶ Ibid.

¹²⁷ Baudrillard J. Op. cit. (The Spirit of Terrorism...). P. 73.

regions of the world¹²⁸ and conditions of lack of information caused, for example, by massive viral diseases¹²⁹ and other social upheavals.

While summarizing the ideas about the influence of the relativization of truth in the context of information influence, it is crucial to note that Baudrillard agreed that manipulations in the information field are a historical characteristic of human civilization¹³⁰. But a change in the information reproduction system, along with a change in socio-political and cultural conditions, entails transformations in our understanding of reality and fiction in the information field. Disinformation, which has become commonplace in the interpretation of postpositivists and postmodernists in the conditions of the 20th century, has been transformed in the 21st century along with the advent of a new generation of media that has accelerated the process of information, and, accordingly, disinformation dissemination. The realities described by postmodernists have become largely prophetic for the new era of digital social media, and the creation and dissemination of the modern “Post-truth” concept (close to postpositivist epistemology and the postmodernist vision of (post)modernity) confirms it.

1.2 Post-truth and modern reflection on the role of truth and emotions in international relations

The appearance in the public discourse of the “Post-truth” concept (first in the English language space and later in other languages) marked a growing interest of the public and the scientific community in the conceptual perception changes of truth and falsehood, fact and fiction in the XXI century. The term immediately received public attention, acquiring supporters of the new concept and its opponents. At the same time, growing interest from politicians, publicists, and scientists left a distorting imprint on it, partially turning it into a rhetorical stamp when used in a generalized way. The specificity

¹²⁸ Francis Z. Bush, *The Media & Misinformation Surrounding the Iraq War* [Electronic resource]. – URL: <https://www.anewseducation.com/post/iraq-war-misinformation> (accessed: 07.02.2023); Moore J. *Bush’s War For Reelection: Iraq, the White House, and the People. Bush’s War For Reelection / J. Moore.* – Hoboken, N.J: Wiley, 2004. – 400 p.; Zemlyanskiy A.V. *Propaganda Tasks of the US Media During the Iraq War / A.V. Zemlyanskiy // Proceedings of Voronezh State University. Series: Philology. Journalism.* – 2022. – № 3. – P. 103-108. (In Russian).

¹²⁹ Yamey G. *Donald Trump: a political determinant of covid-19.* / G. Yamey, G. Gonsalves // *BMJ.* – 2020.

¹³⁰ Baudrillard J. *Op. cit. (The Spirit of Terrorism...).* P. 8.

of post-truth corresponds to the journalistic nature inherent in this term, preventing it from giving an exact definition of the concept and linking it with any particular field of knowledge. For this reason, if we do not consider the notion's stigmatizing properties in public discourse, it has the nature of a new philosophical concept, an existing information threat, and a characteristic of the current state of society that brings postmodernity to a new level.

An active appeal to the idea of the beginning of the post-truth era¹³¹ emerged in speeches, newspapers, and journals released in 2016-2017 and dedicated to events in the West that caused a wide public outcry. These were events considered as unlikely and not accepted as possible by a large part of the expert community: the election of D. Trump as President of the United States in 2016 and the results of the UK European Union membership referendum (the Brexit referendum)¹³². As a result of these two events, confusion arose among the European and American public. It appeared due to the expert knowledge crisis and public trust (in both campaigns) to (dis)information unsuccessfully disproved in the public space¹³³.

Events in the UK and the US have sparked a global discussion about the new status of fake news in world politics. During the two campaigns, there was an active appeal to unverified and unreliable information used for political agitation and argumentation of their positions by politicians. Social networks played a central role in forwarding false information.

An illustrative case was one of the pro-Brexit slogans that the UK is sending £350m a week to the EU instead of investing it in the healthcare system¹³⁴. Despite the refutation of this thesis in public discussions and subsequent campaigning, the power of persuasion and its emotional appeal to the supporters of Brexit prevailed over the arguments of the

¹³¹ Davies W. The Age of Post-Truth Politics [Electronic resource]. – URL: <https://www.nytimes.com/2016/08/24/opinion/campaign-stops/the-age-of-post-truth-politics.html> (accessed: 24.01.2021).

¹³² Freedland J. Op. cit.

¹³³ Gibbs A. 'Post-truth' is the Word of the Year thanks to Brexit, Trump's election campaigns [Electronic resource]. – URL: <https://www.cnbc.com/2016/11/16/post-truth-is-the-word-of-the-year-thanks-to-brexit-trumps-election-campaigns-says-oxford-dictionaries.html> (accessed: 08.02.2023); Rose J. Brexit, Trump, and Post-Truth Politics / J. Rose // Public Integrity. – 2017. – Vol. 19. – № 6. – P. 555-558.

¹³⁴ The Boris Johnson Brexit Bus Lie of £350m [Electronic resource]. – URL: <https://www.conversion-uplift.co.uk/brexit/the-boris-johnson-brexit-bus-lie-of-350m/> (accessed: 08.02.2023); Henley J. Why Vote Leave's £350m weekly EU cost claim is wrong : Politics / J. Henley // The Guardian. – 2016. – 10 Jun.

experts¹³⁵. Along with open public campaigning, where similarly unreliable theses spread directly, targeting attempts were made on social networks to deliver arguments for leaving the EU to potential supporters and undecided voters. Active online campaigning, loaded with unreliable but attractive populist theses, became an unexpected factor for observers and experts, which played a significant role in the voting results. Similar scenarios occurred during the 2016 elections in the United States.

The events in the UK and the United States determined the growing popularity of the term “Post-truth” in the public discourse of Western countries and later in world politics generally. A global acknowledgment of the concept occurred after they chose post-truth as the word of the year by the Oxford Dictionary at the end of 2016¹³⁶. The dictionary defines post-truth as follows: “*Relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief.*”¹³⁷ This interpretation later became referential.

However, the formulation lacks terminological and scientific accuracy. The premise that public opinion formation should a priori happen through an appeal to objectivity indicates this. Given the specifics and complexity of defining the meaning of the objective fact in the public discourse, the presented interpretation has a heuristically significant component, emphasizing the subjectivization of public opinion through the prism of emotions. This component characterizes post-truth as something different from, for example, propaganda because it delegates the status of an active subject in perceiving information to an individual or a social group. The person receiving information determines the degree of truthfulness based on internal motives, e.g., emotional attractiveness or ideological affinity. However, there are no separate determinants in the formulation. They appeared only in the course of further elaborations on the post-truth concept and the conditions associated with it.

It is important to note that the Oxford Dictionary is not the first definition of post-truth. From a historical point of view, post-truth as a separate concept appeared as early

¹³⁵ Op. cit. (British public still believe Vote Leave...).

¹³⁶ Oxford Word of the Year 2016 | Oxford Languages [Electronic resource]. – URL: <https://languages.oup.com/word-of-the-year/2016/> (accessed: 24.01.2021).

¹³⁷ Op. cit. (Oxford Word of the Year 2016).

as 1992 in an article by Serbian-American screenwriter Steve Tesich in *The Nation* magazine¹³⁸. Using the example of American society during the Gulf War, the author points out that “we, as a free people, have freely decided that we want to live in some post-truth world.”¹³⁹ It happened because of the rejection of mechanisms that allow determining the degree of truthfulness of any declared facts due to the unwillingness to undervalue their national self-esteem during the conflict and constantly perceive the negative news agenda¹⁴⁰. In this environment, there is no need to suppress the truth. It is only necessary to hide “inconvenient facts” in favor of those society would perceive with higher enthusiasm. Tesich comes to such conclusions based on an analysis of the official statements of the US ruling circles: “leading to serious military and political consequences, which after some time turned out to be false and have been refuted, but did not have the slightest consequences for those who voiced such intentional “fakes.”¹⁴¹ Tesich’s post-truth model shares characteristics with Baudrillard’s concept of hyperreality, indicating that the real becomes absorbed by desired images, like simulacra disconnected from conventional truth. It is also indicative that Tesich’s article and the later works of Baudrillard¹⁴² lead to similar conclusions when analyzing one event - the Gulf War.

Further development of the term happened in the book of Ralph Case “*The Post-Truth Era: Dishonesty and Deception in Contemporary Life*,”¹⁴³ released in 2004 and considering the problems of lying and its motives. Case suggests that we live in a post-truth era with a growing “ethical twilight zone” that “allows us to dissemble without considering ourselves dishonest.”¹⁴⁴ The growth of the twilight zone is associated with the development of numerous euphemisms that allow hiding lies in a discursive context.

The book contains a list of English words – *imaginative truth*, *alternative reality*, *strategic misrepresentations* – serving, in the author’s opinion, as an indicator of the

¹³⁸ Tesich S. Op. cit.

¹³⁹ Ibid. P.13.

¹⁴⁰ Ibid.

¹⁴¹ Luchinsky Yu.V. “The Government of Lies”: The Phenomenon of Steve Tesich’s “Post-Truth” / Yu.V. Luchinsky // *Information Age*. – 2018. – № 2-1. – P. 323. (In Russian).

¹⁴² Baudrillard J. Op. cit. (*The Spirit of Terrorism...*).

¹⁴³ Keyes R. Op. cit.

¹⁴⁴ Ibid. – P.13.

transition of the speaking subject into the sphere of information misrepresentation. At the same time, any discursive field becomes gradually modified, supplemented by new indicator words. Over the past decade, the idiomatic phrase “highly likely”¹⁴⁵ has become an illustrative example. It was used, in particular, by Western politicians in their attempts to attribute cyberattacks and election interferences.

Significant new publications that popularized “Post-truth” did not appear until a decade after the work of Ralph Case. In the wake of events in the US and UK in 2016-2017, there has been a surge in publications on the topic. In particular, books by such authors as J. Ball¹⁴⁶, M. d’Ancona¹⁴⁷, E. Davis¹⁴⁸, and others¹⁴⁹ have appeared. They have become a kind of intellectual reaction to the public shock from the election of D. Trump and Brexit. The authors responded with their contribution to the declaration of “Post-truth” (and synonyms) as the word of the year in several countries at once¹⁵⁰, perceived as a social indicator of the topic’s relevance. The books revealed, “several important dynamics concerning what exactly ‘post-truth’ politics and ‘fake news’ are, what caused them and what should be done about them.”¹⁵¹ They emphasize the changes in modern politics and the information space.

Despite the appeal to a single issue, the authors, laying the foundation for the study of post-truth, rather multiply interpretations than give an exact definition of the phenomenon. In each of the books, post-truth has a distinct basis of analysis. Davis and Ball are close to the idea of H. Frankfurt, a professor at Princeton University, about the boundaries between lies per se and blatant untruths detached from reality¹⁵². For Davis, post-truth is “a new style of communication replete with attention-grabbing propositions that have no basis in fact or expert judgement at all.”¹⁵³ M. d’Ancona does not make

¹⁴⁵ Hiley Likely, Retirement Novichok. Experts named the words of the year in Russia (In Russian) [Electronic resource]. – URL: https://www.dp.ru/a/2018/11/23/Hajli_lajkli__pensionnij (accessed: 08.02.2023).

¹⁴⁶ Ball J. Op. cit.

¹⁴⁷ D’Ancona M. Op. cit.

¹⁴⁸ Davis E. Op. cit.

¹⁴⁹ Fuller S. Op. cit. (Post-Truth: Knowledge...).

¹⁵⁰ Wight C. Post-Truth, Postmodernism and Alternative Facts / C. Wight // *New Perspectives*. – 2018. – Vol. 26. – № 3. – P. 17-29.

¹⁵¹ Crilley R. International relations in the age of ‘post-truth’ politics / R. Crilley // *International Affairs*. – 2018. – Vol. 94. – № 2. – P. 417.

¹⁵² Frankfurt H.G. *On Bullshit* / H.G. Frankfurt. – Princeton, NJ: Princeton University Press, 2005. – 81 p.

¹⁵³ Davis E. Op. cit. P. XI.

similar distinctions between forms of falsehood. Yet, he singles out public indifference to the devaluation of truth and “the infectious spread of pernicious relativism disguised as legitimate skepticism” as characteristics of post-truth¹⁵⁴. The monographs by Ball, Davis, and d’Ancona became the basis for a new wave of post-truth studies. However, they were marked to a greater extent not by a scientific, but by a journalistic character, expressed in a generalized argumentation.

The transition of post-truth from journalistic to academic space happened only at the end of the 2010s. By this time, the concept had finally entered the political and scientific lexicon. A characteristic feature of post-truth becomes its interdisciplinarity. The term does not develop within the framework of a separate scientific school or concept but gets concurrently through the discourse into many scientific areas. By the beginning of the 2020s, post-truth came into psychology¹⁵⁵, law¹⁵⁶, philosophy¹⁵⁷, media studies¹⁵⁸, medicine¹⁵⁹, and numerous other disciplines. As a result, an interdisciplinary understanding of the phenomenon becomes created, referring to the general problem of the transforming perception of truth and reality in society.

The growth of academic interest in post-truth is visible on the Scopus global database of peer-reviewed scientific literature. The data for 2016-2021 demonstrate a multiple increase in research referring to the post-truth problem. For the entire period, the total number of publications that include the term “Post-truth” in the title or keywords is 1062 documents: articles, books, book chapters, reviews, and reports. At the same time, if in 2016 the number was 14 publications and in 2017 - 147, then, starting from 2018, the annual increase of new publications on the topic is, on average, 225 (see Table 1).

¹⁵⁴ D’Ancona M. Op. cit. P.2.

¹⁵⁵ Effron D.A. Op. cit.; Tsipursky G. A Psychological Approach to Promoting Truth in Politics: The Pro-Truth Pledge / G. Tsipursky, F. Votta, J.A. Mulick // *Journal of Social and Political Psychology*. – 2018. – Vol. 6. – № 2. – P. 271-290.

¹⁵⁶ Op. cit (Post-Truth, Philosophy and Law).

¹⁵⁷ Fuller S. Op. cit. (The Post-Truth About Philosophy and Rhetoric).

¹⁵⁸ Gunkel D.J. Op. cit.; Cosentino G. *Social Media and the Post-Truth World Order: The Global Dynamics of Disinformation. Social Media and the Post-Truth World Order* / G. Cosentino. – Cham: Springer International Publishing, 2020. – 147 p.

¹⁵⁹ Heinrich S. Op. cit.

Table 1 – Number of publications in Scopus with the keyword “post-truth”

Year	Number of publications
2021	222
2020	233
2019	221
2018	225
2017	147
2016	14

Source: compiled by the author based on information from the Scopus database

Another important component of post-truth identified in an in-depth bibliometric analysis¹⁶⁰ is its close connection with concepts such as disinformation (there are two terms in English: misinformation and disinformation) and fakes. A common problem of information influence leads to the interdependence of the concepts and their overlap in the research. At the same time, multiple characteristics of post-truth, such as interdisciplinarity, increased interest after 2016, and a certain Western-centricity, also apply to the aforesaid terms¹⁶¹. Still, it is necessary to draw a line between post-truth and various forms of fake news. Post-truth remains a state of reality in which fakes and disinformation are particularly effective. In other words, if post-truth is a characteristic of an environment, disinformation is a tool of manipulation used within that environment.

After the period of increasing popularity of the term “Post-truth,” another significant frontier in the study of the phenomenon has been the problem of misinformation spread during the COVID-19 pandemic (especially in 2020-2021), also known as “Infodemia” (i.e., the chaotic spread of unverified and false information about the disease by online users)¹⁶². The COVID-19 pandemic demonstrated that fakes and

¹⁶⁰ Patra R.K. Bibliometric analysis of fake news indexed in Web of Science and Scopus (2001-2020) / R.K. Patra, N. Pandey, D. Sudarsan // *Global Knowledge, Memory and Communication*. – 2022.

¹⁶¹ Wang C.-C. Fake News and Related Concepts: Definitions and Recent Research Development / C.-C. Wang // *Contemporary Management Research*. – 2020. – Vol. 16. – P. 145-174.

¹⁶² Infodemic / WHO [Electronic resource]. – URL: <https://www.who.int/health-topics/infodemic> (accessed: 04.12.2022); Akimenko G.V. Infodemy: fake news in a pandemic / G.V. Akimenko [et al.] // *Bulletin of social Sciences and Humanities*. – 2021. – Vol. 2. – № 3. – P. 43-52. (In Russian).

misinformation are universal issues, as they can be linked to political processes just as much as public health problems. The new nature of disinformation illustrated its large-scale potential in the post-truth era and its ability to provoke serious consequences in various social spheres¹⁶³. Infodemia indicated an increased threat regarding the content and speed of content dissemination as much as the possible unintentionality of initial dissemination due to information pressure on information space users¹⁶⁴. Thus, the pandemic brought disinformation and information manipulation beyond the narrow political and electoral segment of the information space, illustrating the growing reality of post-truth with its tendency to emotionality and unverifiability of information, especially in crises.

The pandemic has provided new empirical material for research into the dissemination of falsehoods online. In this regard, there has been an increase in the number of research papers¹⁶⁵ linking the problem of post-truth to the field of medicine and public health¹⁶⁶. It demonstrated the flexibility of the post-truth concept in the study of the information environment. The shifting attention from political issues, which resulted from Brexit and the elections in the United States and other countries, extrapolated the “Post-truth era” idea to the entire media environment, not only political discourse.

In Russian science, “Post-truth’s” entering the scientific discourse coincided with the stages of the concept’s development in the Western community. The appearance of this word in the Russian media similarly aroused interest first among publicists and later among scientists. According to information from the Russian database of scientific articles eLibrary, the total number of studies on the word 'post-truth' in the title or

¹⁶³ Das R. Rethinking Fake News: Disinformation and Ideology during the time of COVID-19 Global Pandemic / R. Das, W. Ahmed // IIM Kozhikode Society & Management Review. – 2022. – Vol. 11. – № 1. – P. 146-159; Kolotaev Y. Op. cit. (Disinformation in Times of Pandemic...).

¹⁶⁴ Zarocostas J. How to fight an infodemic / J. Zarocostas // The Lancet. – 2020. – Vol. 395. – № 10225. – P. 676; The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review / Y.M. Rocha [et al.] // Journal of Public Health. – 2021. – P. 1-10.

¹⁶⁵ Pool J. Infodemic, Misinformation and Disinformation in Pandemics: Scientific Landscape and the Road Ahead for Public Health Informatics Research / J. Pool, F. Fatehi, S. Akhlaghpour // Studies in Health Technology and Informatics. – 2021. – Vol. 281. – Infodemic, Misinformation and Disinformation in Pandemics. – P. 764-768; Martianov V.S. Coronavirus Pandemic and Expert Knowledge Crisis: Reload of Miracle, Mystery and Authority / V.S. Martianov, V.N. Rudenko, L.G. Fishman // Journal of Institutional Studies. – 2022. – Vol. 14. – № 2. – P. 47-58. (In Russian).

¹⁶⁶ Pool J. Op. cit.; Prasad A. Anti-science Misinformation and Conspiracies: COVID-19, Post-truth, and Science & Technology Studies (STS) / A. Prasad // Science, Technology and Society. – 2022. – Vol. 27. – № 1. – P. 88-112.

keywords for 2016-2021 is 542 articles. Since 2018, there has been a steady increase of more than 100 papers per year (see Table 2). At the same time, the central research centers, based on the number of publications, are St. Petersburg State University (67 papers), Lomonosov Moscow State University (34), and the Russian Presidential Academy of National Economy and Public Administration (27).

Table 2 – Number of publications in eLibrary by the keyword “post-truth”

Year	Number of publications
2021	126
2020	120
2019	111
2018	109
2017	74
2016	2

Source: compiled by the author based on information from the electronic library eLibrary

In Russia, similar to the international level, there is no single or dominant approach to the new phenomenon, and interdisciplinarity is its main characteristic. Several significant publications in Russian science worth highlighting contributed greatly to the study of post-truth. A key publication among them is the collective monograph “Post-Truth Politics and Populism,”¹⁶⁷ published under the auspices of St. Petersburg State University. One year before the book’s publication, the author’s team organized one of the first Russian conferences on post-truth, “Post-Truth Politics and Populism in the Modern World,” in 2017¹⁶⁸. Both the book and the conference proceedings indicated the first stage of domestic comprehension of the new phenomenon.

The key feature of this understanding was the connection between post-truth and populism. This connection was due to the necessity of attributing post-truth to a political

¹⁶⁷ Op. cit. (“Post-Truth Politics” and Populism).

¹⁶⁸ Op. cit. (Post-truth politics in the modern world).

phenomenon, allowing to outline the empirical basis of post-truth politics. Yet, the link between post-truth and populism provided a methodological and theoretical basis for analyzing the new phenomenon. As noted in the monograph, “the terms 'post-truth' and 'post-truth politics' remain very vague and are more often used in an ideological and propaganda segment of modern mass media.”¹⁶⁹ The authors, accepting the contradictory and ambiguous nature of the concept, seek to compensate for the lack of conceptual rigor of post-truth at the time by referring to the articulation of “Truth” and “Lies” in political discourse, taking into account the specificities of new media technologies, Web 2.0., and the virtual environment.

Another significant contribution to the Russian understanding of post-truth is the work “Governability and discourse of virtual communities in the conditions of post-truth politics,”¹⁷⁰ published by another team of authors at St. Petersburg State University in 2019. The monograph also considers the conceptual limits of post-truth and its origins. The concept’s duality distinguishes post-truth as a linguistic device and a scientific category, complicating the formation of a scientific understanding of the phenomenon. Skepticism about the very novelty of the idea becomes stressed by the authors, referring to the fact that the main contradictions inherent in the concept are historically universal. At the same time, the paper identifies a list of main characteristics allowing to speak about the specificities of post-truth as a characteristic of the state of modernity. Mediatization, increased emotionality of politics, growing mistrust, polarization, and the uncontrollability of communication are among such specificities¹⁷¹. Together, they define the peculiarity of the historical moment and explain the reason for the ongoing transformations. A distinctive role belongs to “*virtual communities,*” which are the main transformation link and the intermediary of communication and information in the post-truth era.

¹⁶⁹ Op. cit. (Post-truth politics in the modern world). P.6.

¹⁷⁰ Op. cit. (Governability and discourse of virtual communities).

¹⁷¹ Ibid. P.10-19

In recent years, several significant works on strategic communication and transmedia¹⁷², hybrid warfare¹⁷³, and moderation practices¹⁷⁴ addressed the post-truth issues. An important feature of the new studies is their specialization by subject area and their focus on the practical interpretation of their findings.

Despite the interdisciplinarity of the post-truth concept observed in both Russian and foreign sciences, an important point is its connection with *politics and political science*, manifested at the stage of the concept's formation itself. The appeal to post-truth in Western science resulted from the crisis of political institutions and public space, in which dialogue gave way to conflict and polarization. These processes became reflected in electoral campaigns, as well as in interstate relations. The misshaping of previous ideas about political consensus forced the scientific community to return to questions about the principles of forming a common understanding of truth, the characteristics of power to set common meanings, and the ability to transform them. Interpreting this in the category of Foucault's "Discourses of truth," an update in the reflections on the actor and object of narrative power or information influence happened.

A comparison and analysis of currently available political science studies on post-truth reveals the following crucial political characteristics of the phenomenon:

1) The study of political discourse plays an essential role. Post-truth is considered a category with a direct manifestation in the discursive space, where the perception of various political forces is created and falsified¹⁷⁵. Therefore, discourse analysis becomes the key methodology of post-truth research in the applicative direction¹⁷⁶. From the practical and research perspective, a special place in the discursive space belongs to "Strategic narratives" that make politicians and scholars pay attention to the communicated message¹⁷⁷. The narrative creates a politician's self-representation, but it

¹⁷² Ponomarev N.F. Op. cit.

¹⁷³ Ponkin I.V. Op. cit; Popescu N. Hybrid tactics: Russia and the West / N. Popescu // European Union Institute for Security Studies, Alert. – 2015. – Vol. 46.

¹⁷⁴ Models and practices of political content management in the online space of modern states in the era of post-truth / N.A. Ryabchenko [et al.]. - FLINTA, 2021. - 340 p. (In Russian).

¹⁷⁵ Chugrov S.V. Op. cit.

¹⁷⁶ Farkas J. Post-Truth, Fake News and Democracy: Mapping the Politics of Falsehood. Post-Truth, Fake News and Democracy / J. Farkas, J. Schou. – Routledge, 2019. – P. 15.

¹⁷⁷ Michelsen N. Post-Truth-Telling in International Relations / N. Michelsen, B. Tallis // New Perspectives. – 2018. – Vol. 26. – № 3. – P. 12.

does not have to be truthful, only to satisfy political goals, and, thus, semantic uncertainty becomes the system-forming feature of a political discourse based on phantom political meanings¹⁷⁸. The phantoms reinforce the possibility of refuting most political meanings by stigmatizing any of them with discursive clichés (e.g., as fakes), which discredits the value of a narrative regardless of its degree of truthfulness.

2) There is frequent recourse to post- or anti-politics, signaling the end of the previous understanding of the political process¹⁷⁹, in which politics was a competition of fact-based meanings rather than falsifications. Post-truth is a symptom of post-politics, which is “a reaction – a delayed one – to a decades long neoliberal consensus, in which the technocratic management of the economy effectively stifled genuine political contestation.”¹⁸⁰ This reaction embodies “new forms of political experiences, online and offline movements, and a new kind of political consciousness, which does not necessarily follow the logic of political institutions and is sometimes anti-political.”¹⁸¹ Furthermore, “politicians who keep on thinking in directives and exclusive facts cannot respond in a situation where the new populists do not follow the old routines.”¹⁸² The political space is changing, and it is visible through the activities of new actors and the abandonment of old forms of political interaction.

3) H. Arendt’s ideas about “Truth” and politics are fundamental to the political science understanding of post-truth¹⁸³. Two central ideas stand out. The first establishes politics as a sphere in which there is no truth from the outset. “The sphere of real politics is seen by her as 'limited,' where truth is constantly subject to 'destruction.’”¹⁸⁴ By its conflictual nature, power endangers facts¹⁸⁵ by distorting them. The second idea points to

¹⁷⁸ Op. cit. (“Post-Truth Politics” and Populism). P. 34.

¹⁷⁹ Kostyrev A.G. Post-politics in the Nets of Post-truth / A.G. Kostyrev // *Polis. Political Studies*. – 2021. – P. 64-75. (In Russian); Sepczyńska D. Post-Truth from the Perspective of Hannah Arendt’s Political Theory / D. Sepczyńska // *Filozofia*. – 2019. – Vol. 74. – P. 219.

¹⁸⁰ Newman S. Post-truth and the crisis of the political. / S. Newman // *Soft Power*. – 2019. – Vol. 6. – № 2. – P. 99.

¹⁸¹ Browne C. Conceptualising the Political Imaginary: An Introduction to the Special Issue / C. Browne, P. Diehl // *Social Epistemology*. – 2019. – Vol. 33. – P. 393.

¹⁸² Böcskei B. “Post-truth politics” as the normal state of politics / B. Böcskei // *Tamara Journal of Critical Organisation Inquiry*. – 2017. – Vol. 15. – № 3/4. – P. 260.

¹⁸³ Arendt H. *Between Past and Future* / H. Arendt, J. Kohn. – Annotated edition. – New York: Penguin Classics, 2006. – 320 p.

¹⁸⁴ Op. cit. (“Post-Truth Politics” and Populism). P. 12.

¹⁸⁵ Arendt H. Op. cit. P. 220.

the difference between rational and factual truth¹⁸⁶, of which the second depends on the dynamics of the political space, and the first does not apply to politics as such. The rational truth is strictly determined, while the factual truth can have both descriptive and normative scope, i.e., it can become constructed in political debates and discussions¹⁸⁷. Consequently, factual truth is typical for the political space, where an interdependence of fact and opinion exists unacceptable within rational truth. Together with political opinion, it belongs to the same reality – the public sphere¹⁸⁸, in which, however, there is a common understanding of truth¹⁸⁹ expressed in consensus and necessary for the political process.

Attention to the ideas of H. Arendt complements post-truth studies with such variables of the political process as consensus and political opinion. They act as vulnerabilities, the disruption of which can lead to a society of “total lies” that casts doubt on the reality of events.

4) A distinctive feature of post-truth in modern politics is its manipulative nature. Disinformation and hate speech are becoming the main forms of political action for opposing political forces¹⁹⁰. New technologies become manipulation tools, allowing immediate response to information campaigns and direct communication with the audience. The informational power¹⁹¹ concentrates in the hands of digital platforms. They can create algorithms for information distribution or moderation and determine the presence of particular political forces in the information environment¹⁹². At the same time, political content¹⁹³ has become a specific form of information in the digital space, which has become widespread.

¹⁸⁶ Arendt H. Op. cit. P. 220.

¹⁸⁷ The Routledge Handbook of Political Epistemology / eds.. M. Hannon, J. de Ridder. – 1st edition. – London ; New York: Routledge, 2021. – P. 56.

¹⁸⁸ Sepczyńska D. Op. cit. P. 212; Hyvönen A.-E. Careless Speech: Conceptualizing Post-Truth Politics / A.-E. Hyvönen // New Perspectives. – 2018. – Vol. 26. – № 3. – P. 36.

¹⁸⁹ Newman S. Op. cit. P. 26.

¹⁹⁰ Gaultney I. Political polarization, misinformation, and media literacy / I. Gaultney, T. Sherron, C. Boden // Journal of Media Literacy Education. – 2022. – Vol. 14. – № 1. – P. 59-81; Hate Speech and Polarization in Participatory Society / eds. M. Pérez-Escobar, J.M. Noguera-Vivo. – Taylor & Francis, 2022.

¹⁹¹ Seneviratne K. Myth of ‘Free Media’ and Fake News in the Post-Truth Era / K. Seneviratne. – SAGE Publications Pvt Ltd, 2020.

¹⁹² Helberger N. The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power / N. Helberger // Digital Journalism. – 2020. – Vol. 8. – № 6. – P. 842-854; Cobbe J. Algorithmic Censorship by Social Platforms: Power and Resistance / J. Cobbe // Philosophy & Technology. – 2021. – Vol. 34.

¹⁹³ Ryabchenko N.A. Political content management: new linguistic units and social practices / N.A. Ryabchenko, V.V. Katermina, O.P. Malysheva // Church, Communication and Culture. – 2019. – Vol. 4. – № 3. – P. 305-322.; Visvizi A. Politics and Technology in the Post-Truth Era / A. Visvizi, M.D. Lytras. – Emerald Group Publishing, 2019. – 320 p.

5) The loss of unified information dissemination centers in the form of traditional media leads to the involvement of an increasing number of actors in the information generation process. Politicians, political parties, official bodies, public organizations, bloggers, business structures, etc., generate political content in social networks without intermediaries¹⁹⁴. As a result, “all kinds of actors are now seen to be able to influence public opinion through new technologies, leading to competitive escalation and wider disorder.”¹⁹⁵ Such decentralization results in that “the processes occurring at the micro level in social networks and online space generally determine the production and reproduction of political content.”¹⁹⁶ Personalization and orientation on the reactions of individual users have localized the political space, leading to digital exclusion¹⁹⁷ and common dilettantism¹⁹⁸ based on communication abundance and information overload¹⁹⁹. The result is highly personalized political content that forms political echo chambers²⁰⁰, where consumed content corresponds to the user’s worldview.

In sum, three main variables of the post-truth concept’s political understanding reflect the formation of post-truth politics. These include the categories of “*Truth*” (as a universal factual representation of a phenomenon), “*Power*” (as the ability to exercise political will), and “*Consensus*” (as a synthesis of trust and agreement on any socially significant issues). Based on these three variables, the condition for the onset of post-truth politics was the destruction of the existing matrix, in which truth was a universal category, achieved by public consensus and ensured through power mechanisms. In the post-truth era, as a result of consensus disappearance and political polarization, power breaks up into separate discourses of truth, each of which is generated at the individual level, creating a distinct emotional connection.

¹⁹⁴ Giusti S. Op. cit. P. 9-10.

¹⁹⁵ Michelsen N. Op. cit. P. 10.

¹⁹⁶ Ryabchenko N.A. Presidential Campaign in Post-Truth Era: Innovative Digital Technologies of Political Content Management in Social Networks Politics / N.A. Ryabchenko, O.P. Malysheva, A.A. Gnedash // Polis. Political Studies. – 2019. – № 2. – P. 104. (In Russian).

¹⁹⁷ Orlov M.O. The Diversity of the Digital Environment in a Risk Society/ M.O. Orlov // Izvestiya of Saratov University. Philosophy. Psychology. Pedagogy. – 2019. – Vol. 19. – № 2. – P. 159. (In Russian).

¹⁹⁸ Denilkhanov A.K. Op. cit. P. 37.

¹⁹⁹ Baranov N.A. Soft power in the conditions of post-truth / N.A. Baranov // Social and Political Research. – 2018. – № 1. – P. 26. (In Russian).

²⁰⁰ Jamieson K.H. Echo Chamber: Rush Limbaugh and the Conservative Media Establishment. Echo Chamber / K.H. Jamieson, J.N. Cappella. – First Edition. – Oxford; New York: Oxford University Press, 2008. – 320 p.; McIntyre L. Post-Truth / L. McIntyre. – Cambridge, MA: The MIT Press, 2018. – 236 p.

The monograph “A Political Theory of Post-Truth” by I. Kalpokas, a researcher in political theory and international relations, supplements the above variables with two more components that reflect the terminological essence of post-truth. On the one hand, post-truth is “*a general condition of detachment of truth-claims from verifiable facts.*”²⁰¹ On the other hand, such truth-claims are “*pitched to audiences as narrative fictions that constitute their own lived realities and explain the world.*”²⁰²

Kalpokas’s definition has several other advantages. First, it moves away from the opposing dichotomy of truth and emotion²⁰³. Irrationality gives way to “empathy,” helping to understand a particular discourse of truth. The truthfulness of the statements remains relevant, but only to the extent that it corresponds to the audiences’ attitudes. Secondly, the appeal to narrative fiction brings Kalpokas’s point of view closer to postpositivist positions regarding the connection between truth and power. Truth (factual truth) does not refer to reality but only to the possibility of establishing a socially significant interpretation of reality. Thirdly, the formation of separate realities in the discourse indicates trends similar to those identified in the works of Baudrillard and Lyotard. It links post-truth to the mediatization of social reality and the instrumentalization of big data, thereby confirming the technical prerequisites for post-truth creation in the conditions of (post) modernity.

The given generalizations indicate that, in modern world politics, emotional affiliation and not the information’s content becomes a priority. Thus, the issue of emotional perception and its amplification becomes central. This forces political scientists to turn to the psychological mechanisms used to reinforce the truthfulness of a thesis, which is especially important in studying manipulative technologies²⁰⁴ in the form of disinformation or hate speech. In world politics, these problems are of particular interest because such informational manipulations can be used maliciously as part of information influence and propaganda.

²⁰¹ Kalpokas I. A Op. cit. P. 5.

²⁰² Ibid.

²⁰³ Durnová A. Understanding Emotions in Post-Factual Politics / A. Durnová. – Cheltenham: Edward Elgar Publishing, 2019. – 192 p.

²⁰⁴ Rogalsky A. I. Expert discussion about the phenomenon of “post-truth”: domestic and foreign discourse / Rogalsky A. I. // Socio-political processes in a changing world. - 2020. - No. 20. - P. 122. (In Russian).

Therefore, the analytical application of “post-truth” implies both the development of ideas about the blurring of truth’s essence in the discourse and its transformation into the subject of an emotional response exacerbated through new media. As a result, in the political science comprehension of post-truth, the main focus is shifted from the relativism of any discourse of truth to the issues of *emotional influence on the dissemination of data, information, and truth* because of the transition “from a rationality oriented era to an emotionality oriented one.”²⁰⁵ It means recognizing, “both within the study of these respective fields and within wider public discourse, that emotions matter.”²⁰⁶ The discourse of truth, which downplays the importance of emotions in the political space, is being replaced by a post-truth discourse, taking into account and emphasizing the emotional dimension.

The development of “post-truth politics” as an academic concept is, consequently, heuristically useful in analyzing modernity through the lens of political and international studies for multiple reasons. First, it creates a comprehensive characteristic of the current state of the global political discourse and the information production and dissemination process. Secondly, it contributes to the modern information threats study and the understanding of new informational and digital factors of international relations. Thirdly, being inherently heterogeneous and interdisciplinary, post-truth illustrates the continuity of the research agenda from postpositivism and postmodernism, supplementing it with current processes in the technological and social spheres.

1.3 Critique and conceptual limitations of a scientific application of post-truth to the study of information influence in international relations

The theoretical understanding of post-truth and the emergence of the general contours of the concept itself do not suspend the term’s several problems and challenges. Among them, a series of issues of a *general critique* of “Post-truth,” caused by internal contradictions and disagreements on the term by scholars and publicists, as well as a list

²⁰⁵ Rogalsky A. I. Op. cit.

²⁰⁶ Crilley R. Op. cit.

of *methodological challenges*. They include further theoretical and methodological reflection and issues of the practical application of post-truth in the analysis and study of international political processes.

From the point of view of the problems of scientific application and philosophy of science, post-truth has three areas of criticism related to *scientism*, *originality*, and *reproducibility* of knowledge, where scientism appears to be the most fundamental problem. Section 1.2 of this chapter partially addresses some key aspects, highlighting the inherent limitations of post-truth, associated with its publicist nature, political discourse²⁰⁷, and multiple contradictory definitions²⁰⁸ that complicate the study of the phenomenon. Post-truth bears an imprint of a political cliché serving politicians to discredit their opponents²⁰⁹. Calling something “post-truth” means bringing the opponent out of the current discourse of truth, i.e., to delegitimize him in front of a part of the audience.

However, the gradual process of scientific understanding of post-truth has provided it with a substantial component that refers to the conditions of a polarized society in a technogenic reality²¹⁰, where mutual accusations of loss of truthfulness create parallel political hyper-realities. The clarification of the characteristics of post-truth within individual disciplines and the general growth of publications coincide with the deepening of the philosophical and political understanding of post-truth through its connection with earlier ideas and concepts. For this reason, the further development of post-truth, in the presence of constructive criticism, will bring the term (and the ideas behind it) to a new level of understanding and scientific elaboration.

At the moment, the most acute question is the originality of the ideas behind post-truth for international relations since, from the critics’ point of view, lies, emotionality,

²⁰⁷ Tuzovskii I.D. Post-truth as a syndrome of Digital Age: ultimate understanding of the phenomenon and scenarios of the future / I.D. Tuzovskii // *Philosophical Thought*. – 2020. – № 12. – P. 42 - 60. (In Russian); Kolotaev Y. Op. cit. (Prospects and limitations of the scientific use...).

²⁰⁸ Kolotaev Y. Op. cit. (Prospects and limitations of the scientific use...).

²⁰⁹ Habgood-Coote J. Stop Talking About Fake News! / J. Habgood-Coote // *Inquiry: An Interdisciplinary Journal of Philosophy*. – 2019. – Vol. 62. – № 9-10. – P. 1033-1065; Op. cit. (“Post-Truth Politics” and Populism). P. 6; Carrera P. The stratagems of post-truth / Carrera P. // *RLCS. Revista Latina de Comunicación Social*. – 2018. – Vol. 73. – P. 1469-1481; Op. cit. (Governability and discourse of virtual communities). P. 9.

²¹⁰ Salgado S. Online media impact on politics. Views on post-truth politics and postpostmodernism / S. Salgado // *International Journal of Media & Cultural Politics*. – 2018. – Vol. 14. – № 3. – P. 317-331.

and manipulation are not fundamentally new in the 21st century compared to previous eras²¹¹. “Post-truth is a new name for ideological truth,”²¹² while if there is post-truth, there must have been an era of truth²¹³. The same criticisms apply to the linguistic component of the new term and its essential characteristics, revealing the weakest features of the concept.

However, they become the starting point for the conceptualization of post-truth. To prove that “the introduction of the term post-truth into scientific discourse [...] is not only justified but also necessary”²¹⁴ means referring to the preconditions that have shaped post-truth conditions. The valid claim that the problems of lying and emotionalism in politics are historically universal should not deny the presence of dynamics (renewal) of informational and political reality. Renewal represents post-truth’s specificity, combining society’s immanent problems with modern issues.

In the scientific space, three factors explain the decrease in the influence degree of facts in public discourse in the XXI century: *technological breakthrough*, polarization, and trust crisis. The first of them, the technological leap, is caused by the emergence of social media and digital algorithms²¹⁵. In contrast to traditional media, which often have quantitative and qualitative limitations in their information activities, social media pose significant challenges to the old information dissemination system. A key catalyst for these changes has been the emergence of the Internet, once considered a force for democratization but later becoming a source of digital challenges²¹⁶. The decentralization and mediatization²¹⁷ of the information creation process not only brought digital freedoms

²¹¹ Giusti S. Op. cit.; Fuller S. Post-Truth: Knowledge As A Power Game: A Political Theory. Post-truth / S. Fuller; transl. A. Smirnova. – HSE Publishing House, 2021. – 368 p. (In Russian); Op. cit. (The Routledge Handbook of Political Epistemology); Picciotto R. Is evaluation obsolete in a post-truth world? / R. Picciotto // Evaluation and Program Planning. – 2019. – Vol. 73. – P. 88-96; Penzina A.I. Post-truth: Origin and Current Situation / A.I. Penzina // Language and speech on the Internet: Personality, society, communication, culture. - Peoples Friendship University of Russia (RUDN), 2020. – P. 222-228. (In Russian).

²¹² Rostova N.N. Philosophical Analytics of Idea of the Post-Truth / N.N. Rostova // Christian Reading. – 2018. – № 6. – P. 131. (In Russian).

²¹³ Habgood-Coote J. Op. cit.; Rietdijk N. Post-Truth, False Balance and Virtuous Gatekeeping / N. Rietdijk, A. Archer // Virtues, Democracy, and Online Media: Ethical and Epistemic Issues / eds. N. Snow, M.S. Vaccarezza. – Routledge, 2021.

²¹⁴ Tuzovskii I.D. Op. cit.

²¹⁵ Lewandowsky S. Op. cit. P. 359; Farkas J. Op. cit. P. 55.

²¹⁶ Lewandowsky S. Op. cit.

²¹⁷ Kalpokas I. A. Op. cit. P. 14.; Marcinkowski F. Mediatization of Politics: Reflections on the State of The Concept / F. Marcinkowski // Javnost - The Public. – 2014. – Vol. 21. – № 2. – P. 5-22.

but also led to a massive proliferation of fakes on the web, as it allowed any user to replicate their content to any audience without any significant restrictions.

However, the actual decline in the manageability and rationality of the international political media space has occurred through the proliferation of content personalization algorithms. A digital attention economy²¹⁸ and sensation culture²¹⁹ created a digital person cut off from the information space of other members of online communication²²⁰. Such conditions lead to an information culture of iteration (replacing the culture of cognition)²²¹, where the repetition of empathically appealing information grips the user and places him in conditions favorable to political manipulation.

The technological background of post-truth is strengthened by the second component - *political polarization*. The fictitious pluralism and democratization of the Internet have led to social disunity, but its origins lie beyond the digital world. Global inequality, increasing interest in populism²²², and distrust of traditional media²²³ or other social and political institutions resulting from multiple crises (economic, migration, etc.) began to split the global and national public into independent segments - echo chambers²²⁴. A person experiencing social frustration has the opportunity to find supporters in the digital society, forming a closed information field that confirms the correctness of their attitudes.

Post-truth affirms that in world politics, when a conflict situation arises, truth recedes into the background, the factual truth becomes replaced by political opinion, and social contradictions can not be resolved within a rationalist approach²²⁵. The emergence of the new populism in world politics, bringing back sharp social debates to Western

²¹⁸ Harsin J. Regimes of Posttruth, Postpolitics, and Attention Economies / J. Harsin // *Communication, Culture & Critique*. – 2015. – Vol. 8. – № 2. – P. 327-333.

²¹⁹ Chugrov S.V. Op. cit. P. 45; Crilley R. Op. cit. P. 421.

²²⁰ Pariser E. Op. cit.

²²¹ Foster C. Truth as social practice in a digital era: iteration as persuasion / C. Foster // *AI & SOCIETY*. – 2022.

²²² Lewandowsky S. Op. cit. P. 359.; Martinets Y.A. The Growth of Right-Wing Populism in Modern Europe (According to the Estimates of European Experts) / Y.A. Martinets // *RSUH/RGGU Bulletin Series "Political Science. History. International Relations"*. – 2019. – № 1. – P. 123-138. (In Russian).

²²³ Visvizi A. Op. cit. P. 152; *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*. / J.A. Tucker [et al.]. – Rochester, NY, 2018. – P. 53.

²²⁴ Jamieson K.H. Op. cit.

²²⁵ Böcskei B. Op. cit. P. 260.

countries²²⁶, has played a significant role. The rhetorical devices of populism, amplified by digital media, aim to strike at those components of social life that can activate the politically inactive part of the population. Brexit and D. Trump, as well as numerous populist movements and parties in Europe and elsewhere, are examples of how post-truth politics is taking hold through social polarization.

The third component of current conditions is the *growing public distrust* of “dominant narratives” in various spheres of life. It means a decline of trust in expert knowledge²²⁷, the devaluation and crisis of objective facts in the public sphere²²⁸, the general reduction and disappearance of meanings²²⁹ leading to the delegitimization of scientific knowledge²³⁰, as well as the formation of many current examples of hotbeds of tension (in particular, in the context of the COVID-19 pandemic) that create mutual misunderstanding and distrust of society and the authorities²³¹. An erosion of trust in governance institutions hinders the restoration of the status of experts and scientists. At the same time, scientists fail to predict social upheavals and conflicts (Brexit, COVID-19 pandemic, aggravation of international relations), which only contributes to the crisis of trust.

The anti-expertise combined with the lack of unity of the reality retranslation channels leads to the destabilization of reality itself. Such an environment becomes favorable for conspiracies²³² because, in conditions of the destruction of trust in facts,

²²⁶ Podrezov M.V. The History of the Study of Political Populism and Its Current Position in Political Science / M.V. Podrezov // Tomsk State University Journal. – 2020. – № 451. – P. 97. (In Russian). Shmatenko L.M. Russian Disinformation Campaigns: What Risk for the European Union and the International Liberal Order? / L.M. Shmatenko // Academic Workshop The EU and the Crisis of the International Liberal Order: A Systemic Crisis? – 2019; Shibkova M.O. Populism and Euroscepticism: Correlation of Notions / M.O. Shibkova // Contemporary Europe. – 2019. – № 4 (89). – P. 124. (In Russian); Noury A. Identity Politics and Populism in Europe / A. Noury, G. Roland // Annual Review of Political Science. – 2020. – Vol. 23. – № 1. – P. 421-439.

²²⁷ Newman S. Op. cit. P. 95.

²²⁸ Garbuznyak A.Yu. Post Truth Phenomenon: The Devaluation Of Fact In Media Discourse / A. Yu. Garbuznyak // Knowledge. Understanding. Skill. – 2019. – № 1. – P. 184-192. (In Russian).

²²⁹ Koshkarova N.N. On the way to truth leading to lies: the phenomenon of post-truth in modern political communication / N.N. Koshkarova, N.B. Ruzhentseva // Political Linguistics. – 2019. – № 1. – P. 51. (In Russian).

²³⁰ McIntyre L. Op. cit.

²³¹ Eremina N.V. Op. cit.; Tianyang F. Op. cit.

²³² Shatin Y.V. Post-Truth as a Rhetorical Phenomenon in Modern Media Space / Y.V. Shatin // Vestnik NSU. Series: History and Philology. – 2020. – Vol. 19. – № 6. – P. 252; Vinogradova T.V. Science and Technology Research (STS) and Post-Truth Politics. (Review) / T.V. Vinogradova // Social and humanitarian sciences. Domestic and foreign literature. Ser. 8, Science of Science: Abstract Journal. - 2018. - №. 1. - P. 66-79. (In Russian).

“any coherence [...] raises suspicions of manipulation”²³³ among users of digital space. “The mass fabrication of things and events makes people equally open to all possible narratives,”²³⁴ and destructing metanarratives, already identified in postmodern philosophy, are, thus, embodied in digital reality.

Post-truth is, therefore, novel for several reasons. Even relying on ideas originating in postmodernism and postpositivism, post-truth accentuates new challenges of the international system, primarily related to the information space. They relate to the technical changes in the information dissemination medium reinforced by the current political polarization by social media and the crisis of trust combined with the loss of unified narratives. An understanding of post-truth of this kind is a response to the existing skepticism or even concern in academia about the relevance of its application²³⁵. Post-truth forms its own heuristic space (the plurality of truths) with its research questions (connection with emotions), responding to the demands of modernity, even if the phenomenon’s origins lie in human psychology and history.

A final aspect of the critique of post-truth remains the question of its reproducibility and universality. Post-truth studies have, for now, a distinct Western ontology and the empirical base of Western countries. This situation makes the existing view of post-truth largely *Western-centered*²³⁶, which may limit the application of the concept in the global dimension. Remaining particularly typical for the early years of the development of the post-truth concept after 2016, questions of Western democracy maintenance continue to dominate the research agenda around post-truth even now²³⁷. Democracy and freedom often become the central objects onto which post-truth politics becomes projected²³⁸. More acutely presented, this translates into a democratic eschatology²³⁹ unfolding amid

²³³ Ocheretyany K.A. Faket - a unit of digital experience / K.A. Ocheretyany // Philosophical analytics of the digital era. - St. Petersburg: St. Petersburg State University Publishing House, 2020. - P. 315. (In Russian).

²³⁴ Ibid. P. 316.

²³⁵ Khroul V. M. "Post-truth" in a Global Context: Popularity, Heuristic Value and Ontological Essence / V.M. Khroul // Communications. Media. Design. – 2022. – Vol. 7. – № 1. – P. 5-22. (In Russian).

²³⁶ Farkas J. Op. cit.

²³⁷ Giusti S. Op. cit. P. 65.; Lewandowsky S. Op. cit. P. 354.; Farkas J. Op. cit.; Michelsen N. Op. cit. P. 8.; Sim S. Post-Truth, Scepticism & Power / S. Sim. – Cham: Springer International Publishing, 2019. – P. 36.

²³⁸ Farkas J. Op. cit. P. 63.

²³⁹ Ibid.

systemic problems and crises within Western countries, where democracy commonly in Western academia becomes interpreted in a liberal sense.

For this reason, there are concerns that the concept becomes modified or irrelevant when the study of post-truth touches other parts of the world²⁴⁰. However, this formulation of the problem remains controversial because of the question of how unique Western democracy is in the post-truth context. In Russian academia, for example, post-truth is often considered in the context of liberalism but from the perspective of its general decline²⁴¹. Post-truth signals the processes leading to the erosion of Western liberalism from within. In other words, while Western scholars emphasize that democracy needs protection from post-truth, Russian scholars show that the established system of social relations has flaws that have crystallized under post-truth conditions.

Consequently, the question of the mutual influence of liberal democracy and post-truth has a significant position within the whole concept, yet simplifying the discussion of post-truth to questions of democracy is reductionist. The initial interest in threats to Western democracy stimulated attention to post-truth and continues to maintain its relevance for understanding Western society. But post-truth can also have considerable explanatory power in studying non-democratic or non-Western regimes. In recent years, numerous studies have expanded the geography of post-truth studies. It happened in terms of the objects of study (new countries or regions) and the research centers themselves.

The Middle East, where post-truth is associated with notions of authoritarianism rather than democracy, is essential²⁴². The study of authoritarianism as the primary source of post-truth politics, in general, occupies a distinct niche within the overall concept²⁴³. Studies of the Global South²⁴⁴ and the Asian region²⁴⁵ are also gaining significance. As a result, despite the general criticism of post-truth for its Western-centricity, interest in the

²⁴⁰ Block D. Op. cit. P. 123.

²⁴¹ Chugrov S.V. Op. cit. P. 42-59; Denilkanov A.K. Op. cit.; Kolotaev Y.Y. Op. cit. (Russia's identity formation...); Gromyko A.A. New Reality and European Security / A.A. Gromyko // Contemporary Europe. – 2016. – № 5 (71). – P. 5-10. (In Russian).

²⁴² Cosentino G. Op. cit. (Post-truth politics in the Middle East...); Medeiros J. de. Op. cit.; Douai A. Op. cit.

²⁴³ Sim S. Op. cit. P. 50.

²⁴⁴ Swartz L. Op. cit.

²⁴⁵ Yee A. Post-Truth Politics and Fake News in Asia / A. Yee // Global Asia. – 2017. – Vol. 12. – P. 66-71; Tomsa D. Public Opinion Polling and Post-truth Politics in Indonesia / D. Tomsa // Contemporary Southeast Asia. – 2020. – Vol. 42. – № 1. – P. 1-27; Siddiquee Md.A. The portrayal of the Rohingya genocide and refugee crisis in the age of post-truth politics / Md.A. Siddiquee // Asian Journal of Comparative Politics. – 2020. – Vol. 5. – № 2. – P. 89-103.

concept and the phenomenon behind it is generally growing in the social and technical sciences and international relations, in particular. Important is that post-truth, through gradual scientific conceptualization, could develop a set of characteristics (technical and social preconditions) that are universal to the modern world. Therefore, post-truth as an analytical concept is reproducible in different political and cultural contexts, considering the specificities of a given community, country, or region.

Overall, the criticism of the post-truth concept highlights that, despite the incomplete solution of all problematic aspects, post-truth withstands the critics in terms of the concept's academic application. The transformation of modern reality requires the introduction of new, clarifying definitions that allow us to search for relevant ways to overcome pressing problems and challenges. Thus, the conceptualization of post-truth as a state of politics and society, and, at the same time, as an actual threat in international relations, is a way to overcome the contradictions surrounding the term. Post-truth is a system-wide challenge, provoked not so much by individual actors as by the aggregate international environment.

The theoretical understanding of the problems of the information factor in modern international relations and "Truth" in the information society led to the following conclusions. The formation of a new information environment, viewed through the prism of postpositivism, postmodernism, and post-truth, indicates the dual nature of the truth displacement from its dominant position in society. On the one hand, they are embedded within the classical philosophical question of the role and significance of truth and falsehood in human society. On the other hand, the philosophical schools and concepts mentioned in this study illustrate the transformation of ideas about the role of facts and emotions in contemporary politics under the influence of ideological, social, and technological changes and threats.

The postpositivist understanding of the intersubjective nature of truth allows us to express it as a social construct, and the notions of "Hyperreality" and "Simulacra"

illustrate the current conditions of truth devaluation in the context of informatization and the spread of mass media. In the modern political environment, this creates a space favorable to the influence of information, capable of reproducing events and phenomena that have no connection with reality. Such 21st-century conditions manifest themselves in post-truth politics, which summarizes the philosophical understanding of information influence and combines technical transformations with the current social crises of polarization and loss of trust.

In political studies, post-truth appears in the scope of destruction of the truth, power, and consensus concepts, where universal representations become the object of interpretations generated by multiple actors at the individual level. Despite the conceptual difficulties, post-truth has secured its heuristic value, expressed in ideological interdisciplinarity, emphasis on technogenic transformations, attention to emotions as an attribute of political and international interaction, and the study of the process of spreading lies in the digital society.

A crucial goal in the further evolution of the post-truth concept is to expand its field of application beyond the problematization of the phenomenon in liberal democratic regimes and institutions. The rejection of normativity is essential for its application in the study of information conflicts and the development of universal means to reduce the impact of post-truth on social development.

Chapter 2: The structure of digital information influence in international relations in the post-truth era

The post-truth era, with all its key features, has become an environment of social polarization, interethnic and interstate confrontation, and, above all, information wars. In such conditions, the conflicting parties aim to discredit or informally influence the opponent or the public. The main tools are information weapons in the form of disinformation, fake news, strategic communication, and hate speech. They exist in the information environment, while the realities of the 21st century bring it to the digital environment associated with the development of the information society²⁴⁶, in which information is both the primary good or source of production and a potential threat when aimed at distortion or deception.

The study of information manipulation issues related to post-truth requires a systematic mapping of the principles of application of “Post-truth” in the analysis of contemporary international and political processes. It means paying attention to the *logic of information influence*, including the aspects of the manipulation process and the connection between them, *influence tools* and digital threats (disinformation and hate speech), as well as *resources and measures to combat* digital threats.

Existing models of information influence in media and world politics have different variables to analyze. In particular, some emphasize the agent, message, and interpreter²⁴⁷, or identify the main stages of the information distribution process, such as message creation, production, and dissemination²⁴⁸. This dissertation, however, focuses on three actors of the information influence process. These include the source or subject of information influence, the information dissemination platform or intermediary, and the object of information influence. The subject and the object are of primary importance, as a semantic link appears between them mediated by the platform. The role of platforms

²⁴⁶ Webster F. *Theories of the Information Society* / F. Webster. – 4th edition. – Abingdon, Oxon: Routledge, 2014. – 404 p.

²⁴⁷ Wardle C. *Information disorder: Toward an interdisciplinary framework for research and policymaking* / C. Wardle, H. Derakhshan. – Council of Europe report DGI (2017)09. – 2017; Jia F. *Misinformation Literature Review: Definitions, Taxonomy, and Models* / Jia F. // *International Journal of Social Science and Education Research*. – 2020. – Vol. 3. – № 12. – P. 85-90.

²⁴⁸ *Ibid.*

and their meaning is determined by the context of the digital manipulation tools application.

Such a model of interaction takes into account and revises the methodological basis and idea of H. Lasswell's formula²⁴⁹ with its critical additions²⁵⁰, which distinguishes five basic (communicator, message, medium, recipient, effect) and two additional aspects or issues of communication (motives and context). However, in contrast to the linear communication model, the main focus of the study is not on communication per se but on the behavior, strategies, and principles of actors' responses in the creation – dissemination – and receiving of information. For this reason, the focus is on *the communicator, the medium, and the recipient*. Message and effect, as well as, motives and context, are treated as subsidiary issues. They are considered more generally in the modeling process and when reviewing information influence tools.

2.1. Methodology for studying the structure of information influence: individual and social aspects

The creation of a comprehensive representation and, consequently, the development of a method for studying information influence, *which refers to the impact on an individual or society of intentionally or unintentionally disseminated (dis)information*, requires a clear representation of the main input parameters as well as the characteristics of the actors. It is necessary to take into account the parallel globalization of the information network and the localization of the political space, i.e., to analyze and model both the characteristics of society as a whole and the characteristics of the mechanisms that determine the actions of individuals and users of the global network.

Information influence is interpreted in this work mainly in a negative connotation, as a process of spreading false and unreliable information. It is related to the motivation of the subject of information influence, the tool (content), and the result (effect) of

²⁴⁹ Lasswell H.D. The structure and function of communication in society / Lasswell H.D. // The communication of idea / ed. L. Bryson. – New York: Harper and Row, 1948. – P. 37-51.

²⁵⁰ Braddock R. Op. cit.

influence. At the same time, since post-truth is part of the sphere of perceiving lies²⁵¹ and not producing them, studies of information influence must pay special attention to *behavioral and psychological (cognitive) aspects of information processing* at the stage of its creation and consumption. Information influence has a dual nature: it can be purely politicized (deliberately used in confrontation), and it can also be generated without any political subtext (due to uncertainty). It points to the complexity of information influence, supplemented by the notions of object and subject of information manipulation.

In the case of the information influence subject, its main characteristics are the presence or absence of intent and the motivation behind the disseminated information. In international relations, the subject with political intent is key, as it can generate information campaigns and strategic communications that destabilize the established discourse of truth. The act of information influence contains an initial mindset, that is important for the subject, as it reflects his interest. Consequently, the choice of information dissemination strategy will aim at maximizing personal benefit and usefulness. This approach to understanding the subject's behavior is characteristic of *rational choice theory*²⁵², which is why the behavior of information influence subjects is often analyzed through the prism of game theory models²⁵³.

To explain the subject's actions, it is functional to distinguish two forms of relationship with factual information²⁵⁴. The first, in which the generally accepted truth corresponds to the actor's tasks, implies that actions are focused on the intensified dissemination of information to establish a general agreement on its reality. The second form is associated with a situation when the information contradicts the existing

²⁵¹ Wight C. Op. cit.

²⁵² Obadã R. Sharing Fake News about Brands on Social Media: a New Conceptual Model Based on Flow Theory / R. Obadã // Argumentum. Journal of the Seminar of Discursive Logic, Argumentation Theory and Rhetoric. – 2019. – Vol. 17. – P. 159; Tsebelis G. Nested Games: Rational Choice in Comparative Politics. Nested Games / G. Tsebelis. – University of California Press, 1990. – 300 p.; Griggs S. Rational Choice in Public Policy: The Theory in Critical Perspective / S. Griggs // Handbook of Public Policy Analysis. – Routledge, 2007; Kiser E. The Revival of Narrative in Historical Sociology: What Rational Choice Theory can Contribute / E. Kiser // Politics & Society. – 1996. – Vol. 24. – № 3. – P. 249-271.

²⁵³ A Survey of Game Theoretic Approaches to Modelling Decision-Making in Information Warfare Scenarios / K. Merrick [et al.] // Future Internet. – 2016. – Vol. 8. – № 3. – P. 34; Guo Z. Effect of Disinformation Propagation on Opinion Dynamics: A Game Theoretic Approach / Z. Guo, J. Valinejad, J.-H. Cho // IEEE Transactions on Network Science and Engineering. – 2022. – Vol. 9. – № 5. – P. 3775-3790; Deception, identity, and security: the game theory of sybil attacks / W. Casey [et al.] // Communications of the ACM. – 2018. – Vol. 62. – № 1. – P. 85-93.

²⁵⁴ Toward a Comprehensive Model of Fake News: A New Approach to Examine the Creation and Sharing of False Information / A.P. Weiss [et al.] // Societies. – 2021. – Vol. 11. – № 3. – P. 7.

consensus. In such a situation, the actor is forced to ensure the realization of his intentions by disseminating, for example, disinformation²⁵⁵. Consequently, the subject's behavior depends on the correlation between the information he provides and the actual reality.

For the information influence subject in the conditions of post-truth politics, distortion or reinterpretation of facts becomes the dominant behavioral strategy, as the rationalistic need to follow the dominant narrative is lost. The subject acts as an attacker against the existing power and discursive system. In terms of actor modeling, such a subject has four basic behavior models to revise the status quo²⁵⁶:

- *Degradation* – “flooding a victim with non-sensical, redundant or irrelevant data to hide actual facts”²⁵⁷;
- *Corruption* - creating false beliefs by spreading disinformation or replacing true information with false data²⁵⁸;
- *Denial* – “to prevent users from accessing true information by the way of inhibiting true information flow,”²⁵⁹ manifested through technological interference;
- *Subversion* - is based on deceiving users by altering the processing of input data²⁶⁰, the attack targets information processing algorithms and their distortion.

The aggregate of these models is called the Borden-Kopp model²⁶¹, which was brought to the analysis of fake news and disinformation from the game-theoretic modeling of information wars. According to this model, all four types of rational behavior aim at changing the perception of the object of information manipulation and can occur both at the hardware-technical level of information systems and the cognitive level. At the same time, degradation and corruption are more directly focused on the individual, while denial and subversion are applied more easily on the platform level.

As a result of all these models, depending on the initial goal, an actor can achieve a variety of politically significant outcomes: modify consensus, influence public opinion,

²⁵⁵ Op. cit. (Toward a Comprehensive Model of Fake News...).

²⁵⁶ Kopp C. Op. cit.

²⁵⁷ Kopp C. Op. cit. P. 4.

²⁵⁸ Guo Z. Op. cit. P. 3781.

²⁵⁹ Ibid.

²⁶⁰ Guo Z. Op. cit. P. 3781

²⁶¹ Kopp C. Op. cit.

devalue political discourse, delegitimize the media, influence electoral behavior, achieve defamation, polarization within internal or external conflicts, contribute to international tensions, and so on²⁶². Still, to achieve them, an actor needs a comprehensive understanding of the object of information influence, its incentives, and the response mechanisms that determine the success of information influence.

In actor modeling, the object of information influence is organized differently from the subject. It contains both causes and mechanisms of manipulation, as well as a complex system of motives and beliefs that can be modified through manipulation. The object more often reflects the logic of *bounded rationality*²⁶³ since logical judgments about the surrounding world are coupled with emotional perceptions that, together with natural cognitive mechanisms and behavioral stereotypes, illustrate the complexity of an individual's (or social group's) response to information and digital influence. The idea of bounded rationality, proposed by economist H. Simon²⁶⁴, in the early 1980s, emphasizes that people have “very narrow capacities for simultaneous attention to different pieces of information.”²⁶⁵ In this sense, “people are, at best, rational in terms of what they are aware of, and they can be aware of only tiny, disjointed facets of real.”²⁶⁶

Consequently, the irrationality of behavior is primarily conditioned by the limit of human perception to the diversity of information, so the limit of human rationality should be sought, for example, not in emotions but in a person's internal psychological and cognitive mechanisms²⁶⁷. Such mechanisms include the two types of thinking identified by the economist and social psychologist D. Kahneman²⁶⁸. The combination of fast (reflexive or stereotyped thinking) and slow (requiring greater cognitive and analytical costs) thinking explains the systems most vulnerable to information influence. In

²⁶² Giusti S. Op. cit. P. 10.

²⁶³ Nerino V. Overcome the fragmentation in online propaganda literature: the role of cultural and cognitive sociology / V. Nerino // *Frontiers of Sociology*. – 2023. – Vol. 8. – P. 1-10; Op. cit. (Falling for fake news...); Op. cit. (Information overload for (bounded) rational agents); *Routledge Handbook of Bounded Rationality* / Ed. R. Viale. – 1st edition. – London; New York, NY: Routledge, 2020. – 664 p.

²⁶⁴ Arias-Maldonado M. A Genealogy for Post-Truth Democracies: Philosophy, Affects, Technology / M. Arias-Maldonado // *Communication & Society*. – 2020. – P. 68; Simon H.A. Op. cit. P. 301.

²⁶⁵ Simon H.A. Op. cit. P. 301.

²⁶⁶ *Ibid.*: P. 302

²⁶⁷ Durnová A. Unpacking emotional contexts of post-truth. / A. Durnová // *Critical Policy Studies*. – 2019. – Vol. 13. – № 4. – P. 448-449.

²⁶⁸ Kahneman D. *Thinking, Fast and Slow* / D. Kahneman. – 1st edition. – London: Macmillan, 2011. – 512 p.

particular, the swiftness and low cost of the first type of thinking leads people to believe more easily in stereotypical information that does not require them to process it for long periods.

In mass consciousness, for example, people trust generally accepted information because it increases its likelihood to be true²⁶⁹. For this reason, intuitive and reflective adherence to majority opinion is common in public discourse, based on how many other people are committed to a particular view. Similar reasoning, expressing one of many heuristics or empirical rules²⁷⁰, plays a significant role in carrying out crucial life processes by simplifying them. However, this form of thinking plays a negative role and becomes a behavioral trap²⁷¹ when it comes to information that requires triggering the second thinking mechanism. The ability to manipulate the user through the fast thinking is thus one of the essential mechanisms used in information campaigns on the Internet and beyond.

All manipulative techniques rely on cognitive biases (heuristics, behavioral traps)²⁷², representing a wide range of “systematic deviations from rationality in decision-making processes, the main source of which is rooted in the very structure of cognitive processes.”²⁷³ Their immanence in human consciousness makes them a crucial mechanism of manipulation. In a digital society, the most significant aspect of cognitive biases is human exposure to excessive information and the need to react quickly to it. Digital systems that accelerate political processes lead to a flow of information facilitative to manipulation, while the available time and cognitive resources of the object of influence are often insufficient for rational behavior or response²⁷⁴.

²⁶⁹ McDermott R. Psychological Underpinnings of Post-Truth in Political Beliefs / R. McDermott // PS: Political Science & Politics. – 2019. – Vol. 52. – № 2. – P. 220.

²⁷⁰ Halpern D. The Psychology of critical thinking / D. Halpern. - St. Petersburg: Peter, 2000. - 512 p. (In Russian).

²⁷¹ Halpern D. Op. cit.; Plous S. The Psychology of Judgment and Decision Making / S. Plous. - Moscow: Filin, 1998. (In Russian).

²⁷² Kahneman D. Op. cit. (On the reality of cognitive illusions); Tversky A. Judgment under Uncertainty: Heuristics and Biases / A. Tversky, D. Kahneman // Science. – 1974. – Vol. 185. – № 4157. – P. 1124-1131; Cognitive Biases in Visualizations / ed. G. Ellis. – Cham: Springer International Publishing, 2018. – P. 2-3.

²⁷³ Loginov N.I. Current Trends in International Research on Cognitive Biases in Decision-Making Processes / N.I. Loginov, A.I. Aleksandrova // Psychology. Journal of the Higher School of Economics. – 2020. – Vol. 17. – № 3. – P. 445. (In Russian).

²⁷⁴ Op. cit. (Information overload for (bounded) rational agents).

Such circumstances seriously contribute to misinformation through the *availability cascade* widely used in political campaigns. “In this cognitive bias, the reinforcement of collective belief in a piece of information is directly related to the amount of repetition, accessibility, persuasiveness, argumentation, and references to authoritative sources.”²⁷⁵ The digital environment reinforces each of these parameters by facilitating the reproducibility of any information, which directly affects the effectiveness of such manipulation.

Understanding the limited rationality of humans and their inherent heuristics suggests that information influence, like post-truth, “appeals to [cognitive] reflexes but not to reflection.”²⁷⁶ This state of affairs does not mean there is no way to counteract cognitive biases. The awareness of this problem itself promotes approaches related to debiasing strategies. They rely on a critical attitude towards cognitive processes, i.e., an inherent understanding of the bounded rationality of human beings. An example of such an approach is structured analytic techniques²⁷⁷, based on “externalizing and decomposing the cognitive process will result in bias mitigation.”²⁷⁸ Psychological interventions aimed at individual processing of the distortion are also possible²⁷⁹. At the same time, the issue of eliminating biases remains a matter of personal initiative since the awareness of the presence of cognitive distortions is no guarantee of practical action by the individual to mitigate them.

Even though cognitive biases are one of the main mechanisms of various information influence strategies²⁸⁰, the influence object has an additional level of analysis – social²⁸¹. This can mean either the existence of two independent analysis units – a person or society/social group – or two levels of perception within a person subject to influence

²⁷⁵ Mikheev E.A. Disinformation about the new coronavirus infection COVID-19: psychological mechanisms of influence / E.A. Mikheev // Problems of the National Strategy. - 2021. - No. 1 (64). – P. 167. (In Russian).

²⁷⁶ Chugrov S.V. Op. cit. P. 46.

²⁷⁷ Heuer R.J. Structured Analytic Techniques for Intelligence Analysis / R.J. Heuer, R.H. Pherson. – 1st edition. – Washington, DC: CQ Press, 2010. – 368 p.

²⁷⁸ Belton I. Cognitive biases and debiasing in intelligence analysis / I. Belton, M. Dhimi // ed. R. Viale. – London: Routledge, 2020. – P. 552.

²⁷⁹ Ibid. P.553-555

²⁸⁰ Wanasika I. Deception as Strategy: Context and Dynamics / I. Wanasika, T. Adler // Journal of Managerial Issues. – 2011. – Vol. 23. – № 3. – P. 376.

²⁸¹ Calero Valdez A. Studying Biases in Visualization Research: Framework and Methods / A. Calero Valdez, M. Ziefle, M. Sedlmair // Cognitive Biases in Visualizations / ed. G. Ellis. – Cham: Springer International Publishing, 2018. – P. 17.

– individual and social. A two-level understanding of the object means that the influence on an individual occurs in parallel with the influence on society as a whole, whereas the influence on a person activates both individual and social triggers.

The social level of the object builds upon mass consciousness and common cultural and discursive practices. In contrast to the individual level within the social space, psychological mechanisms play a supporting role. The *consensus* on shared ideas about the veracity of major social narratives plays a key role in the public consciousness. Differences in positions between political actors are acceptable for the system as long as they do not violate the underlying consensus. However, as post-truth shows, consensus destruction entails an increase in the irrationality of actors in the sense that they no longer conform to the universalist understanding of social truth.

From the standpoint of poststructuralism, this aspect can be expressed through the extrapolation of discourses of truth to the relationship of conflicting social groups. Each group creates its interpretation of truth to the best of its power abilities. In such a situation, the task of the manipulation subject is to modify the discourse of truth or to create a parallel discourse of truth. It reflects the rationality that interprets truth and facts in the interest of the manipulation subject. In this case, the process of interpretation structures upon the intermediary function of the media transmitting the necessary discourse of truth. Unlike the individual level, it is not cognitive mechanisms that operate here, but media effects that directly impact the whole society²⁸². These media effects include framing, priming, and agenda-setting.

All three types of media influence increase the possibilities of political and discursive power through the media²⁸³ and operate on the principle of active strategies of the influence subject, whether it is “assembling a narrative [...] to promote a particular interpretation” (framing)²⁸⁴, increasing the “raise the salience or apparent importance of certain ideas” (priming), or purposefully “defining problems worthy of public and

²⁸² Scheufele D.A. Op. cit. (Framing, agenda setting, and priming...). P. 10.; Kolotaev Y. Op. cit. (Political Influence of Online Platforms...).

²⁸³ Scheufele D. Op. cit. (Framing As a Theory of Media Effects); Pan Z. Op. cit.; Brewer P.R. Op. cit.

²⁸⁴ Entman R.M. Framing Bias: Media in the Distribution of Power / R.M. Entman // Journal of Communication. – 2007. – Vol. 57. – № 1. – P. 164.

government attention.”²⁸⁵ Through a series of such discursive practices, the context of a particular message, event, or statement can be altered following a particular agenda or discourse of truth. It emphasizes not only the discursive power of the subject but also the potential threat to that power from the media, which takes on the intermediary role. Through the lens of critical discourse analysis, the relationship between politics and the media has led to mediatized politics, which is “sometimes seen as the colonization of politics by the media, and sometimes seen as the colonization of the media by politics.”²⁸⁶

However, such impact means influence mainly on the rational-factual level of social relations because it focuses on completeness and content accuracy issues of the information provided. At the same time, social emotions and sentiments are also an area of direct information influence, which is especially important in a post-truth context, shifting attention from facts to emotions²⁸⁷. In this regard, there is a need to turn to an understanding of emotions as part of the social space²⁸⁸ that characterizes emotions as a manifestation of overlearned cognitive patterns²⁸⁹. Like heuristics, emotion is a repetitive pattern of behavior, but one that is linked not only to individual sensory perception but also to collective influence. This approach, called the “history of emotions,”²⁹⁰ views emotions through their construction by culture and society.

Acts of emotional expression – emotives – are a combination of private and public factors. It gives the process of expressing emotions both descriptive and performative features because “they have the capacity to enact change on the speaking subject or addressee,”²⁹¹ and not just displaying individual feelings. The social dimension of emotion explains why emotional speech acts (such as verbalized fear, aggression, etc.) transcend individual experience. “It comes to matter politically when it is articulated by collectives, usually towards a shared objective.”²⁹²

²⁸⁵ Entman R.M. Op. cit. P. 164.

²⁸⁶ Fairclough N. *Media Discourse* / N. Fairclough. – 1st edition. – London: Bloomsbury Academic, 1995. – P. 200.

²⁸⁷ Kolotaev Y. Op. cit. (*Political Implications of Hate Speech Digitalization...*); Kolotaev Y. Op. cit. (*Post-truth through the lens...*).

²⁸⁸ *Early Modern Emotions: An Introduction*. *Early Modern Emotions* / Ed. S. Broomhall. – London: Routledge, 2016. – 424 p.

²⁸⁹ McGrath L.S. Op. cit.; Süselbeck J. *Sprache und emotionales Gedächtnis. Zur Konstruktion von Gefühlen und Erinnerungen in der Literatur und den Medien* / J. Süselbeck // *Emotionen*. – 2019. – P. 282.

²⁹⁰ Reddy W.M. Op. cit.; McGrath L.S. Op. cit.

²⁹¹ Colwell T.M. *I.2 Emotives and emotional regimes* / T.M. Colwell // *Early Modern Emotions*. – Routledge, 2016. – P. 7.

²⁹² Wahl-Jorgensen K. Op. cit. P. 768.

Such a vision of the nature of emotions led to the emergence of the concept of “Emotional regimes,” which express the dominant models of emotional perception in society²⁹³. In an ideological sense, they are comparable to discourses of truth applied to factual information. The interactive nature of emotions makes them dynamic. As a result, they become “discursively constituted in individuals and are similarly constructed as emotional “regimes” in societies under particular historical conditions.”²⁹⁴ The principles of the emotional regimes’ formation draw on the norms, customs, and political institutions that make emotions themselves governable²⁹⁵ and, therefore, interpretable.

From a political point of view, emotional regimes are the dominant modes of acceptable emotional expression²⁹⁶ imposed by powerful actors. It means that under certain circumstances, the collective understanding and expression of individual emotions (fear, anger, hatred) can be changed under the influence of internal or external factors by the information influence subjects. Such actions threaten social stability since they can be directed to politically significant emotions (anger and hatred)²⁹⁷, which have a serious motivational potential.

In a changing emotional regime, individual social targets of information manipulation (politically marginalized groups, radical movements, etc.) may tend to accept and express such political emotions as a normal practice. This may be particularly acute in the case of hate speech, where the shifts in emotional regimes create the image that there is acceptable hate, which can lead to direct or digital violence. Thus, the information influence subject aims to influence individual social groups through the emotional regimes associated with them.

The ideas of emotional regimes operating in society and the possibility of their change indicate the fundamental vulnerability of the social level and the emotional public sphere²⁹⁸ of the information influence object. In modern realities, this vulnerability

²⁹³ Reddy W.M. Op. cit.

²⁹⁴ Nye R.A. Review of *The Navigation of Feeling: A Framework for the History of Emotions*, William M. Reddy / R.A. Nye // *The Journal of Modern History*. – 2003. – Vol. 75. – № 4. – P. 921.

²⁹⁵ Colwell T.M. Op. cit. P. 7.

²⁹⁶ Garrido S. Emotional Regimes Reflected in a Popular Ballad: Perspectives on Gender, Love and Protest in ‘Scarborough Fair’ / S. Garrido, J. Davidson // *Musicology Australia*. – 2016. – Vol. 38. – № 1. – P. 65.

²⁹⁷ Lyman P. The Domestication of Anger: The Use and Abuse of Anger in Politics / P. Lyman // *European Journal of Social Theory*. – 2004. – Vol. 7. – № 2. – P. 133.

²⁹⁸ Martyanov D.S. Op. cit.

manifests clearly under the influence of algorithmic digital media²⁹⁹, “subjective and often irrational in their representations,”³⁰⁰ just as discourses of truth fall under the influence of media effects.

Such interdependence of the two aspects of the social essence of the object suggests that “the discourse of truth is as dependent on the current emotional regime as the regime itself is on the discourse of truth.”³⁰¹ However, despite this connection, the digital age has emphasized that the actual impact of information is not so much on the whole society but on single users, absorbing individual experience with social variables. Consequently, the digital crowd effect forms as a cumulative interaction of individual users reacting to events through their emotionally and factually personalized information flow.

All of the above representations, including the rational strategies of the subject, bounded rationality, and the duality of the social nature of the object, express the basic analytical categories within the framework of the methodology for studying information manipulation. Their structural relationship is illustrated in Figure 1. The subject of information influence, which can be any social unit from the state to the individual³⁰², acts as the primary unit of analysis with its inherent will and motivation. The connection of the subject to his motivation (personal motives, group, or national interests), expressed in the final *intention* to disseminate information, is the backbone for information influence³⁰³. Since, in most scenarios, information influence involves an attitude important for the subject expressing his interest, the analysis of the subject’s behavior requires a rationalist paradigm and attention to purposeful behavior patterns.

²⁹⁹ Boler M. The affective politics of the “post-truth” era: Feeling rules and networked subjectivity / M. Boler, E. Davis // *Emotion, Space and Society*. – 2018. – Vol. 27. – P. 82-84.

³⁰⁰ Op. cit. (“Post-Truth Politics” and Populism). P. 9.

³⁰¹ Kolotaev Y. Op. cit. (Political Implications of Hate Speech Digitalization...). P.521.

³⁰² Op. cit. (A Survey of Game Theoretic Approaches...). P. 4.; Robinson M. Cyber warfare: Issues and challenges / M. Robinson, K. Jones, H. Janicke // *Computers & Security*. – 2015. – Vol. 49. – P. 70-94.

³⁰³ Ibid.

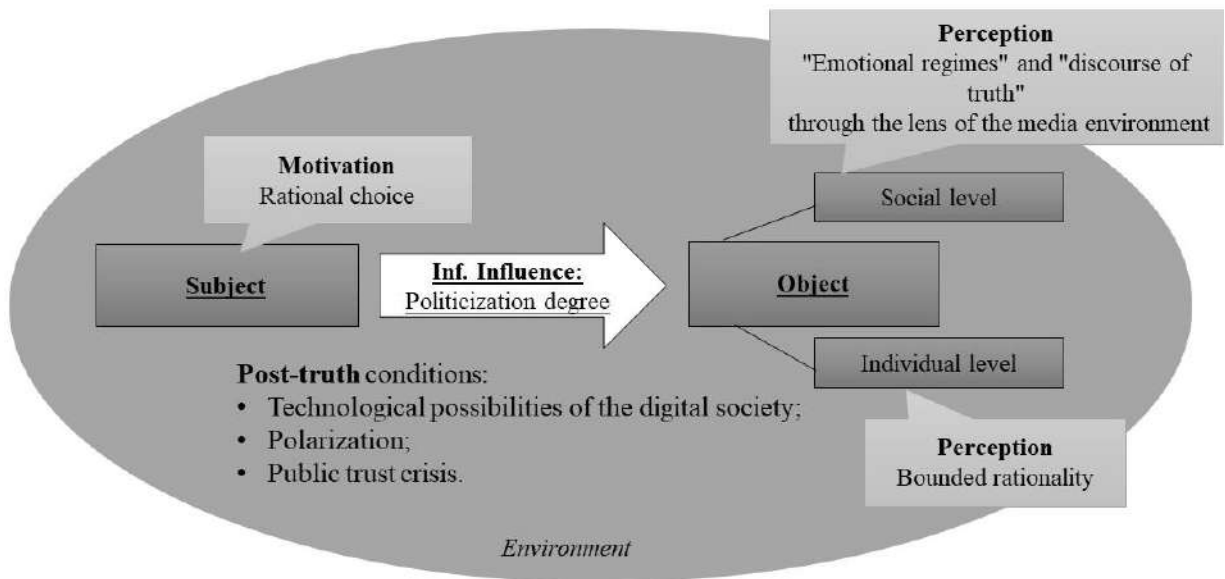


Figure 1 – The relationship of the subject, object, and environment of information influence in post-truth conditions

Source: compiled by the author

The object of information manipulation is a more complex structure. It can be both an individual and a society, and the influence can affect individual and social response mechanisms. In contrast to the subject, creating a clear institutional framework for the object is more difficult because both levels (societal and individual) are closely interconnected. The bounded rationality logic describes most comprehensively the results of information influence at the individual level, indicating that a person has a system of fast but not always rational cognitive mechanisms for processing information, which form cognitive biases and behavioral traps exploited by the subject of information influence.

At the social level, discourses of truth and emotional regimes function as systems for responding to incoming information, responsible for the collective consensus on the veracity of facts and the admissibility of emotions. The close relationship between facts and emotions in the context of information influence is related to the emotional response to the disseminated truth and the facts' influence on the framework of acceptable emotionality. In information manipulation, emotional regimes and discourses of truth hold back disinformation and hate speech, functioning only as long as these mechanisms are not distorted by more serious information influence.

Another characteristic of information influence complements the structure shown in Figure 1, connected to the relationship between the subject and the object. Considering the variety of possible subject characteristics (political/non-political subject; person/automated system, etc.)³⁰⁴, there are diverse sequences of influence, expressed in the fact that subjects can be primary or secondary (Figure 2). It means that one subject acts as the message's primary source, while the other subject might support it. The presence of two subjects for a message does not imply that their intentions and motivations are similar. For example, a secondary subject may initially be an object (recipient) of information influence, whereby he decides to disseminate this information further. The result is an iterative process with a multiplicative consequence.

At the same time, if the primary subject had political goals, the secondary actor could act based on the need to inform other users (friends, relatives, etc.) without realizing the initially manipulative nature of the primary message. This model of behavior builds the repetitive essence of the repetition culture in the post-truth era³⁰⁵, which contributes to the spread of false information in international relations.

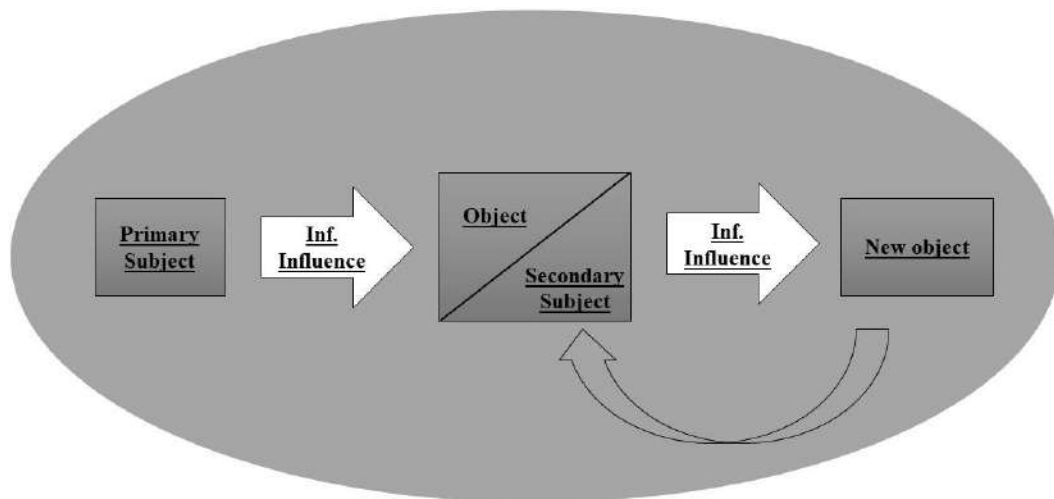


Figure 2 – Primary and secondary objects of information influence. Iterative principle

Source: compiled by the author

Summarizing the above methodological framework for studying information influence in the realities of a digital society, one should emphasize the relationship between the subject and the object, which affects the further creation of intermediate links

³⁰⁴ Giusti S. Op. cit. P. 10.

³⁰⁵ Foster C. Op. cit.

and tools of information influence. Each has its internal behavioral logic, important to mention when comparing both types of subjects.

The depicted information influence model helps to structure the main variables and the nature of manipulative campaigns in the post-truth and digital society. The presented system of analysis contains several intentional simplifications. In particular, at the individual level, the distinction between information influence in the rational and emotional dimensions is deliberately omitted because the irrationality of the object's behavior in the presented system is not caused by reliance on emotional triggers, but by cognitive mechanisms that simplify human behavior and interpret emotions as cognitive patterns. At the same time, as noted above, emotions do not always conflict with rational behavior and reliance on facts. As a result, individual emotionality remains outside the scope of the study because it requires additional consideration of internal mental processes that poorly reflect the political essence of the studied subject.

Thus, within the framework of this study, the bounded rationality of the information manipulation objects, as well as the social essence of emotions and truth manifested in their changeability under the influence of international and social processes, are considered the initial provisions of the information influence practice. In addition, these research frameworks are object-oriented, as they focus less on the motives of the subject and more on the object's reaction. It allows for a further focus on potential responses to incoming information influence and methods for countering it.

2.2. Digital disinformation, hate speech, and the role of online platforms in international processes

The subject's action strategy and the structure of the object's vulnerability reflect the main aspects of information influence. They show the relationship between the main elements of manipulative influence in the discursive or emotional space. However, in the relationship structure between the subject and the object (Figure 1), two additional elements play a significant role: *the intermediary* and *the instrument* of information influence. Intermediaries are those information spaces and platforms through which

information exchange happens between the subject and the object. The instruments are various types of information attacks and manipulations in the information space, the most important of which in this study are *disinformation* and *hate speech*.

The intermediary transmits the message that the subject uses as part of its information activity. The nature of intermediaries was briefly touched upon when considering the media effects of traditional media. In modern conditions, intermediaries include not only the mass media but also a variety of new media: digital platforms, social networks, video hosting, blogs, etc.³⁰⁶ Despite their intermediary function, they not only extrapolate media effects to the digital space³⁰⁷, but they begin to show signs of actorness³⁰⁸. It is most evident in the ability of platforms to act as censors or gatekeepers³⁰⁹ in the transfer of information (Figure 3).

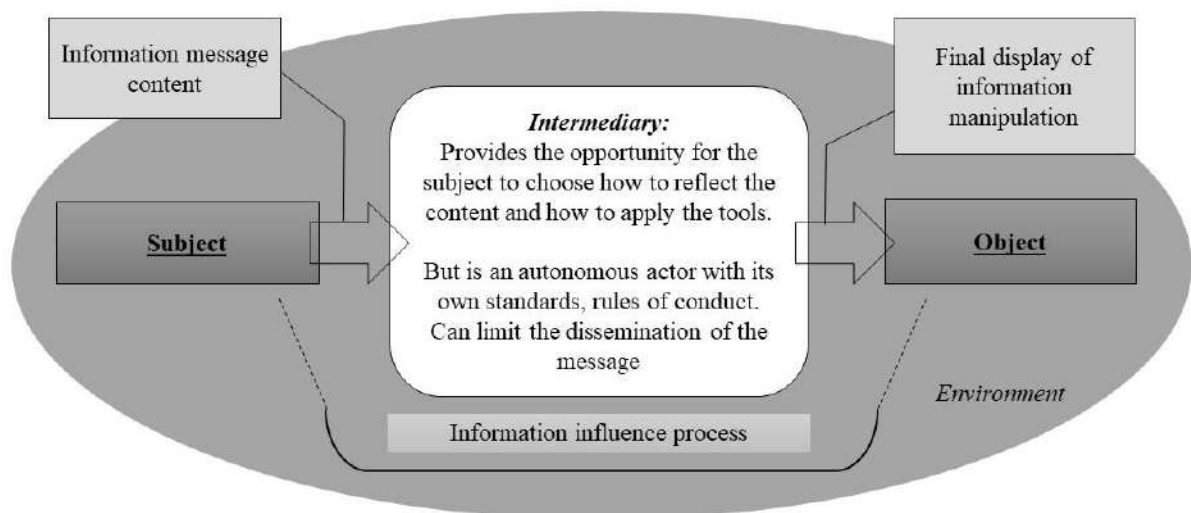


Figure 3 – The role of the intermediary in the information influence

Source: compiled by the author

Social media's interactive and networking capabilities have demonstrated the changing approach to media in the 21st century. The new logic of social media has led to changes in media content production, the principles of information dissemination, and the

³⁰⁶ Rogaleva O.S. New media: evolution of the concept / O.S. Rogaleva, T.V. Shkaiderova // Herald of Omsk University. – 2015. – № 1 (75). – P. 222-225. (In Russian); Nosovets S.G. New media: defining the notion / S.G. Nosovets // Communication Studies. – 2016. – № 4 (10). – P. 39-47. (In Russian); Plantin J.-C. Digital media infrastructures: pipes, platforms, and politics / J.-C. Plantin, A. Punathambekar // Media, Culture & Society. – 2019. – № 2 (41). – P. 163-174.

³⁰⁷ Zhuravskaya E. Political Effects of the Internet and Social Media / E. Zhuravskaya, M. Petrova, R. Enikolopov // Annual Review of Economics. – 2020. – Vol. 12. – № 1. – P. 415-438.

³⁰⁸ Eilders C. Media as political actors? Issue focusing and selective emphasis in the German quality press / C. Eilders // German Politics. – 2000. – Vol. 9. – № 3. – P. 181-206.

³⁰⁹ Nielsen R.K. News media, search engines and social networking SITES as varieties of online gatekeepers / R.K. Nielsen // Rethinking Journalism Again. – Routledge, 2016.

way people use media³¹⁰. The digital mediatization of society has created conditions for political actors to “adapt their discourse to the digital age,”³¹¹ leading to the overall digitalization of world politics and international relations.

The increasing influence of social media on political ideas dissemination has made political discourse dependent on digital platforms: they are used to transfer crucial information and become the arena for manipulation and hate speech. Platforms and social networks determine how users’ messages are reflected³¹², thus influencing the qualitative characteristics of information (whether through the message’s number of characters, the content visibility rules, or moderation) and turning the platform into a communication participant.

Changes in the way users disseminate and receive politically relevant information have shaped the academic community’s perception of social media’s political power³¹³, expressed through active participation and mediation in political and international relations³¹⁴. B.I. Page, a media actorness researcher, emphasizes that the idea of a political actor “implies observable action that is purposive (though perhaps functional rather than consciously intended) and sufficiently unified.”³¹⁵ The mediating function of social networks partly illustrates this.

A typical form of active intervention in the information dissemination process is moderation (algorithmic and non-algorithmic)³¹⁶, with all variations ranging from

³¹⁰ Klinger U. *Network Media Logic: Some Conceptual Considerations* / U. Klinger, J. Svensson. – 2015. – *Network Media Logic*. – P. 34.

³¹¹ Estellés M. *The educational implications of populism, emotions and digital hate speech: A dialogue with scholars from Canada, Chile, Spain, the UK, and the US* / M. Estellés, J. Castellví // *Sustainability (Switzerland)*. – 2020. – Vol. 12. – № 15. – P. 5.

³¹² Baines D. *Defining misinformation, disinformation and malinformation: An urgent need for clarity during the COVID-19 infodemic* : Discussion Papers / D. Baines, R.J.R. Elliott. – Department of Economics, University of Birmingham, 2020. – P. 8.

³¹³ Shirky C. *The Political Power of Social Media: Technology, the Public Sphere, and Political Change* / C. Shirky // *Foreign Affairs*. – 2011. – Vol. 90. – № 1. – P. 28-41.

³¹⁴ Kolotaev Y. *Op. cit. (Political Influence of Online Platforms...)*.

³¹⁵ Page B.I. *The Mass Media as Political Actors* / B.I. Page // *PS: Political Science & Politics*. – 1996. – Vol. 29. – № 1. – P. 20.

³¹⁶ Gorwa R. *Algorithmic content moderation: Technical and political challenges in the automation of platform governance* / R. Gorwa, R. Binns, C. Katzenbach // *Big Data & Society*. – 2020. – Vol. 7. – № 1. – P. 2053951719897945; Dias Oliva T. *Content Moderation Technologies: Applying Human Rights Standards to Protect Freedom of Expression* / T. Dias Oliva // *Human Rights Law Review*. – 2020. – Vol. 20. – № 4. – P. 607-640; Gillespie T. *Content moderation, AI, and the question of scale* / T. Gillespie // *Big Data & Society*. – 2020. – Vol. 7. – № 2. – P. 2053951720943234.

flagging content³¹⁷ to *deleting* or even *deplatforming* users³¹⁸. Through moderation, platforms often seek to demonstrate their social responsibility, showing that they protect the users' interests. Such intentions are represented in community rules that define the platforms' moderation process.

At the same time, soft moderation in the form of information flagging has explicit characteristics of priming and framing, as it increases the visibility of an issue/point of view or, on the contrary, aims to silence it. More severe forms of moderation – removing content or users from platforms (deplatforming) – can have even greater resonance as they demonstrate the platform's ability to alter the agenda. The situation with the removal of former US President D. Trump's Twitter account in 2021³¹⁹ made this notably evident. Such situations directly illustrate how an intermediary defines permissible rhetoric and discourse on its platform while setting a common agenda.

Most of these forms of moderation allow us to talk about the agency and influence of mainstream digital platforms within information mediation because of their mass scale. Unlike traditional media, social networks bring together billions of users. They act as an intermediary between the subject and the object of information influence and, at the same time, have a resource of simultaneous access to massive amounts of data about the users and their preferences, which gives them the ability to target content to the object³²⁰. Along with moderation, content selection algorithms determine the information that reaches the user³²¹. At the same time, targeting and algorithms have a clear commercial orientation

³¹⁷ Chipidza W. The effectiveness of flagging content belonging to prominent individuals: The case of Donald Trump on Twitter / W. Chipidza, J. (Kevin) Yan // Journal of the Association for Information Science and Technology. – 2022. – Vol. 73. – № 11. – P. 1641-1658.

³¹⁸ Rogers R. Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media / R. Rogers // European Journal of Communication. – 2020. – Vol. 35. – № 3. – P. 213-229; Evaluating the Effectiveness of Deplatforming as a Moderation Strategy on Twitter / S. Jhaver [et al.] // Proceedings of the ACM on Human-Computer Interaction. – 2021. – Vol. 5. – CSCW2. – P. 381:1-381:30.

³¹⁹ Fontevecchia A. The Fallacy Of 'Deplatforming' Donald Trump / Forbes [Electronic resource]. – URL: <https://www.forbes.com/sites/afontevecchia/2021/01/20/the-fallacy-of-deplatforming-donald-trump/?sh=b1dcfa411418> (accessed: 10.02.2023).

³²⁰ Knoll J. Advertising in social media: a review of empirical evidence / J. Knoll // International Journal of Advertising. – 2016. – Vol. 35. – № 2. – P. 266-300; Barbu O. Advertising, Microtargeting and Social Media: International Conference on Communication and Education in Knowledge Society / O. Barbu // Procedia - Social and Behavioral Sciences. – 2014. – Vol. 163. – P. 44-49.

³²¹ Social Media Algorithms: Why You See What You See / Georgetown Law Technology Review [Electronic resource]. – URL: <https://georgetownlawtechreview.org/social-media-algorithms-why-you-see-what-you-see/GLTR-12-2017/> (accessed: 10.02.2023).

related to maximizing user retention on the platform for content consumption³²². As a result, digital platforms become dominant players, setting the rules of information transfer, which is necessary to consider when building and analyzing information influence models.

However, not only platforms abuse their perception management capabilities. Information influence subjects use their algorithms to distort users' perceptions. As a result, online platforms indirectly shape the environment in which it becomes possible to spread the main information threats that are *instruments/tools of information influence*. While the online platforms' activity is often commercially oriented, the actions of other information actors aim at political discourse and global processes.

In global politics, the two most significant information manipulation tools (for the subject) are *disinformation* and *hate speech*. They are the threats (to the object) most closely associated with post-truth politics, as they are manipulative and stem from misleading information dissemination. Even though both threats disseminate through a common intermediary, they have different levels and mechanisms of influence:

- Disinformation aims to distort rational perception. It exploits cognitive vulnerabilities by distorting and falsifying information to create the illusion of its credibility and truthfulness.
- Hate speech spreads at the level of emotional perception. It can also involve factual information, but more often it relies on stereotypical thinking, a sense of solidarity and shared identity, and has a targeted group or person against whom a shared sense of hatred is formed.

While the spread of disinformation can take place at the individual level, hate speech is predominantly applied at the group level, as the feeling of hatred requires interactivity. Additionally, disinformation is based on denying or distorting the truth, while hate speech uses one of the most powerful political emotions to discredit an object of hate through consolidation. These generalizations only highlight the general

³²² Bhargava V.R. Ethics of the Attention Economy: The Problem of Social Media Addiction / V.R. Bhargava, M. Velasquez // Business Ethics Quarterly. – 2021. – Vol. 31. – № 3. – P. 321-359; Monetizing disinformation in the attention economy: The case of genetically modified organisms (GMOs) / C.D. Ryan [et al.] // European Management Journal. – 2020. – Vol. 38. – № 1. – P. 7-18.

characteristics of these manipulation tools. In terms of form, methods, and terminology, they are characterized by a greater diversity, which should be considered further.

Disinformation

Within the contemporary conceptualization of “deliberate deception in communication scenarios,”³²³ there are several key concepts, which, depending on the context, include “Disinformation,” “Fake news,” “Propaganda,” and, to some extent, “Strategic communication.”³²⁴ By choosing a specific terminological apparatus, researchers define different semantic emphases, starting from one of these definitions.

Fake news (or fake) is the most relevant formulation of the problem of information manipulation through public information “containing unreliable and unverified information that does not correspond to real facts...”³²⁵ This concept became popular in parallel with post-truth in the context of the 2016 events, which is why the terms are sometimes part of the terminological explanation of each other³²⁶. At the same time, the term “Fake news” is an evident rhetorical stamp and stigma in public discourse³²⁷, and it is excessively associated with the “news” component, which prevents it from being used in a broader context.

“Propaganda” refers to a more comprehensive statement of the problem, as it does not inherently express the malicious nature of the actions but rather represents the process of active and purposeful dissemination of any information³²⁸. It is one of the earliest terms used to summarize the process of information influence. “Strategic communication” also describes a centralized process of communicating specific attitudes and positions to a certain public³²⁹ but with a strict causality of actions and the overall intent of an information campaign.

³²³ Alexander J. Disinformation: A Taxonomy / J. Alexander, J. Smith // IEEE Security & Privacy. – 2011. – Vol. 9. – № 1. – P. 58.

³²⁴ Kolotaev Y.Y. Op. cit. (Fake news, disinformation and strategic communications...).

³²⁵ Ilchenko S.N. Op. cit. P. 12.

³²⁶ Corner J. Fake news, post-truth and media-political change / J. Corner // Media, Culture & Society. – 2017. – Vol. 39. – № 7. – P. 1100-1107; Grech V. Fake news and post-truth pronouncements in general and in early human development / V. Grech // Early Human Development. – 2017. – Vol. 115. – P. 118-120.

³²⁷ Tandoc E.C. Op. cit.

³²⁸ Ellul J. Op. cit.

³²⁹ Defining Strategic Communication / K. Hallahan [et al.] // International Journal of Strategic Communication. – 2007. – Vol. 1. – № 1. – P. 3-35; Holtzhausen D. Strategic Communication: Opportunities and Challenges of the Research Area / D.

As a result, terminological pluralism exists in the studies of information influence problems. However, in the general list of terms, the *political aspects* of information manipulation are most reflected in “Disinformation.” This term emphasizes the process of providing information and, especially, the *distortions*. In the English-speaking world, “Disinformation” (in Russian, “дезинформация”) has two interpretations: misinformation and disinformation, the difference between which is the absence or presence of intention in the process of providing false information. In the conditions of the 21st century, the notion of disinformation is actively spread in the discourse of Western regulators, which signals the actualization of the concept in the practical dimension.

Despite the differentiation of concepts, which defines the limits of the applied categories, it is necessary to note the complementarity of the ideas behind them, referring to specific dimensions of disinformation and explaining their properties. Thus, among the origins of the political science understanding of propaganda in the 20th century, there are several important aspects reflecting the essence of intentional information influence. H. Lasswell, who wrote one of the first works on propaganda in the interwar period, refers to the image of “the war of ideas against ideas,”³³⁰ in which the intentional dissemination and repetition of information help to overcome psychological resistance through the standardization of civil consciousness by media and news. At the same time, propaganda has an object-subject model, where the object is any entity from an individual to a society, and the “strategy of propaganda [...] can readily be described in the language of stimulus-response.”³³¹

Another important representative of propaganda research, J. Ellul, notes the importance of “psychological manipulation”³³² for the mobilizing potential of information influence. He identifies its multiple gradations, among which the category of social propaganda (as opposed to political propaganda) and emotional or irrational

Holtzhausen, A. Zerfass // The Routledge Handbook of Strategic Communication. – Routledge, 2014; Gavra D.P. Category of Strategic Communication: Modern Understanding and Basic Characteristics / D.P. Gavra // Information Age. - 2015. - No. 3. (In Russian).

³³⁰ Lasswell H.D. Op. cit. (Propaganda technique in the World War). P. 12.

³³¹ Lasswell H.D. Op. cit. (The Theory of Political Propaganda). P. 630.

³³² Ellul J. Op. cit. P. 61.

propaganda (as opposed to rational propaganda)³³³ are of particular importance, where the former implies influence on general social and cultural norms (emotional regimes) and the latter aims to resonate with the emotional reaction of the object.

H. Lasswell and J. Ellul equally confirm that propaganda and manipulation are an integral part of political existence and political communication, emphasizing its special significance not so much in totalitarian but precisely in democratic regimes. It is due to their dependence on public opinion, which forces political actors to resort to information influence to gain support, leading in part to a state of “dictatorship of palaver.”³³⁴ This formulation of the problem dating back a century illustrates that the essential aspects behind digital disinformation are not only and mostly related to the actual state of society but to the fundamental interdependence of politics, public opinion, and the media.

Disinformation, while essentially aimed at misleading propaganda, is characterized by the presence of several other important variables: a noise source (appearing at the stage of transmission through an intermediary and creating a distorted view of the data) and an unintended recipient³³⁵ (i.e., an object that has become so despite its desire). The first variable serves as a condition for the distortion of the original information or as a prerequisite to creating false data. The second variable implies that the final recipient of the untruth message becomes misinformed against his will.

At the same time, disinformation has multiple manifestations in world politics. Even though some studies point to the irrelevancy of any differentiation of disinformation³³⁶, in practice, many characteristics allow us to speak about its different types, depending on the motive (ideological, psychological, revenue) and verifiability of information³³⁷, (un)reliability and (un)authenticity of data³³⁸, and the degree of truthfulness of information³³⁹. As a result, the list of phenomena that fall under the

³³³ Ellul J. Op. cit. P. 61-87.

³³⁴ Lasswell H.D. Op. cit. (The Theory of Political Propaganda). P. 631.

³³⁵ Alexander J. Op. cit. P. 58.

³³⁶ Kirchner C. Digital Disinformation: Taxonomy, Impact, Mitigation, and Regulation (Dagstuhl Seminar 21402) / C. Kirchner, F. Roesner // Dagstuhl Reports. – 2022. – Vol. 11. – № 9. – P. 28-44.

³³⁷ Op. cit. (A systematic literature review on disinformation...).

³³⁸ Lemieux V. Leveraging Archival Theory to Develop A Taxonomy of Online Disinformation / V. Lemieux, T.D. Smith // 2018 IEEE International Conference on Big Data (Big Data). – 2018. – P. 4420-4426.

³³⁹ “Fake News” Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content / M.D. Molina [et al.] // American Behavioral Scientist. – 2021. – Vol. 65. – № 2. – P. 180-212.

category of disinformation and fakes multiplies, generating entire taxonomies of interrelated manifestations of false information.

Thus, taxonomies can be distinguished based on: *behavioral aspects* (deception, fakery, lack of information, etc.)³⁴⁰, *different goals* (entertainment and attention, monetization, discrediting, etc.)³⁴¹, and *instrumental embodiment* (clickbait, rumor, trolling, fabrications, etc.)³⁴². From an instrumental point of view, fake news also functions as a subtype of disinformation, which receives the most attention in the study of disinformation due to its status as a buzzword and politicized term.

Russian philologists also offer several significant classifications defining the types of disinformation-forming fakery³⁴³. Doctor of philology S.T. Zolyan designates three definitions of falsification³⁴⁴ based on the truth types subject to falsification: correspondent – inconsistency with reality; coherent – inconsistency or contradiction with a certain discourse; pragmatic – inconsistency with practice and what is beneficial to the subject. This classification not only reflects the problem of the plurality of the truth but also shows how exactly and for what reasons an aspect of reality can be subject to distortion and replacement.

The classifications presented above have many overlapping aspects and allow us to select the necessary taxonomic relationships depending on the goals of research or practice. However, for the information influence system modeled in the first section of this chapter, the division of “дезинформация” by *intent* (a1 or a2 in Figure 4) remains fundamental, similar to the English division between disinformation and misinformation. The distinction of intent is unified by the gap between the truthfulness of data and the truthfulness of the message³⁴⁵. In other words, if a message strives to present its content to be truthful, then the data encapsulated in the message itself is not true (b1). It is equally relevant for disinformation and misinformation.

³⁴⁰ Dalkir K. Op. cit. P. 39.

³⁴¹ Romanova E.R., Op. cit.

³⁴² Op. cit. (A systematic literature review on disinformation...). P. 17.

³⁴³ Ilchenko S.N. Op. cit. P. 6-7.

³⁴⁴ Zolyan S.T. “Fakes” – People Or Texts? Round Table Discussion / S.T. Zolyan // Terra Linguistica. – 2021. – Vol. 12. – № 2. – P. 7-32. (In Russian).

³⁴⁵ Baines D. Op. cit. P. 11-12.

But in the structure of information messages, another form of intentional disinformation is possible, in which both the message and the data themselves claim to be true or synchronously deny it (*malinformation*)³⁴⁶. This situation is related to “truth reconfiguration” (b2) when, for example, an actual fragment of reality is presented as evidence to prove the alleged truth of a claim but there is no real link between them except communicatively. Such a situation is possible when the chronicle of some events is exploited to dramatize or interpret other events (such examples accompany any socially significant news today, especially concerning social and political crises³⁴⁷, natural disasters³⁴⁸, etc.). If this is done unintentionally, the result will be close to the combination of a1 and b1, which can be interpreted, for example, as ignorance or lack of context, and the appeal to truth or untruth will not be of fundamental importance.

Thus, based on the given variables, it is possible to model 3 forms of disinformation (Figure 4):

1. a1 + b1 = intentional disinformation with lies;
2. a1 + b2 = intentional disinformation exploiting the truth;
3. a2 + b1 or a2 + b2 = unintentional misinformation.

³⁴⁶ Baines D. Op. cit. P. 11-12.

³⁴⁷ Truth or Fake - No, this video doesn't show thousands of Syrian migrants trying to cross into the EU // France 24 [Electronic resource]. – URL: <https://www.france24.com/en/tv-shows/truth-or-fake/20220923-no-this-video-doesn-t-show-thousands-of-syrian-migrants-trying-to-cross-into-the-eu> (accessed: 25.05.2023); Fake videos spread online after Chadian president's death // The Observers - France 24 [Electronic resource]. – URL: <https://observers.france24.com/en/africa/20210422-fake-videos-spread-online-after-chadian-president-s-death> (accessed: 25.05.2023).

³⁴⁸ FAKE ALERT: Old video from Maharashtra viral as 'massive cyclone' in Kerala // The Times of India [Electronic resource]. – URL: <https://timesofindia.indiatimes.com/times-fact-check/news/fake-alert-old-video-from-maharashtra-viral-as-massive-cyclone-in-kerala/articleshow/82973398.cms> (accessed: 25.05.2023).

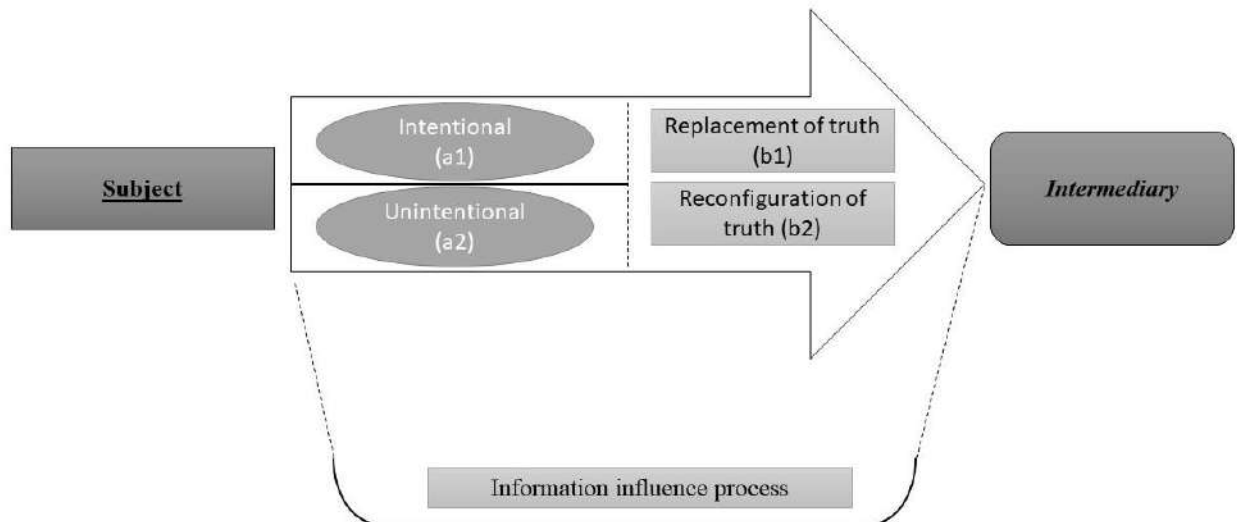


Figure 4 – Forms of disinformation in information influence

Source: compiled by the author

All these three forms of disinformation have a similar effect on the object (they can manipulate its perception), however, they differ in the interactions between the subject and the mediator. The question of the presence or absence of intention directly affects the practices of counteraction, since intention is an important variable in the choice of means of moderation and interaction with the subject. Moreover, the reconfiguration of truth also requires a special approach. It outlines more complex processes of interaction between lies and truth since the communicative reliance on an actual event requires the verification not of the reality of a fact but of the relationship between the message and the presented data.

The types of disinformation reflected in the formulas can resemble, to a certain extent, most of the aforementioned classifications, simplifying them to two variables, which is functionally simpler from the point of view of regulatory practices. This approach does not negate the need for further detailing of various forms of negative information influence. Its task is to systematize multiple types of disinformation within the matrix of relationships between the subject's intention and the truthfulness of the message.

Hate speech

Unlike the widely discussed disinformation in its various manifestations, hate speech is a narrower information threat because it is not directly related to lies and does not aim to undermine the credibility of any information. Its focus is on *provoking a particular emotion in a human community or broadening the acceptance of hateful discourse in a social environment*.

There is no established definition of hate speech in Russian, as the term is a translation of the English concept.

The notion “arose in Western countries as a result of the realization of the need to limit dangerous rhetoric calling for the (continued) exploitation (and in extreme cases, physical elimination) of certain population groups and minorities.”³⁴⁹ Due to different translations of the English original, several words refer to the phenomenon: hate rhetoric, hate speech, etc. (in Russian, “риторика ненависти,” “язык вражды,” “хейтспич,” “язык ненависти”). At the same time, hatred is in some cases replaced by the concept of “hostility,” understood as a *complex affective-cognitive personality trait*³⁵⁰ that manifests itself “in emotions of anger, disgust, and contempt, with their associated feelings and expressions, which can lead to aggressive behavior.”³⁵¹

Conceptual pluralism is also characteristic of hate speech in the English-speaking world, which includes terms associated with the nature of digital content (aggressive, toxic, offensive, hateful content, or cyberbullying)³⁵² or referring to specific reasons for discrimination and hatred, namely political, historical, class, racial or religious differences (racist, anti-Semitic, anti-feminist, xenophobic, Islamophobic, etc.)³⁵³. The complexity of defining and interpreting hate speech is also because the thematic spectrum of hate is constantly dynamic and expanding.

³⁴⁹ Alexeyev A.B. Hate speech and manipulation as peculiarities of the modern political communication / A.B. Alexeyev // Izvestiya of Saratov University. Sociology. Politology. – 2021. – Vol. 21. – № 4. – P. 458. (In Russian).

³⁵⁰ Vasilenko E.N. "Hate Speech": Toward a Definition of a Term / E.N. Vasilenko // Collection of articles of the International Scientific Conference dedicated to the 105th anniversary of Moscow State University named after A.A. Kuleshova. – Mogilev State University named after A.A. Kuleshova, 2019. – P. 127. (In Russian).

³⁵¹ Ilyin E.P. Psychology of communication and interpersonal relations / E.P. Ilyin. - St. Petersburg: Peter, 2009. - P. 254. (In Russian).

³⁵² Offensive, aggressive, and hate speech analysis: From data-centric to human-centered approach / J. Kocoń [et al.] // Information Processing & Management. – 2021. – Vol. 58. – № 5.

³⁵³ Paz M.A. Op. cit. P. 2; Op. cit. (Internet, social media and online hate speech...).

From the point of view of socio-political embodiment, hate speech should be understood as a *public manifestation through linguistic means of hostility or hatred against a specific social group that shares common characteristics, taking the form of insult or incitement to violence and aggression*. At the same time, an important feature of hate speech in international relations is that it often refers to violent acts and can border on extremism, closely related to hate crimes³⁵⁴. Therefore, unlike disinformation, hate speech is more likely to be associated with illegal actions, as it is linked to crimes of racism, chauvinism, discrimination, and harassment disseminated in the information space.

Based on its thematic framework, hate speech forms an independent discursive space with its inherent discourse of truth and emotional regime. In world politics, it is often associated with group resentments or grievances³⁵⁵, stereotypes or prejudices about the object³⁵⁶, and can be expressed through specific linguistic means. Such lexical means tend to delineate group or national affiliation through the “us”/“others” system, resorting to varying degrees of aggressiveness of the communication itself³⁵⁷, for example, through overt insults or positive rhetoric with an implicit negative connotation³⁵⁸. In such a context, lexical units are used to form degrading comparisons, slander, threats, and other forms of hate speech expression³⁵⁹.

These forms of hate speech can be summarized in a “Pyramid of hate,”³⁶⁰ which shows the sequence in the degree of aggression and activity in expressing hostility (Figure 5). As a result of the progression from the lower levels of the pyramid to higher levels, hate speech intensifies until it turns into direct offenses and crimes. Such the pyramid reflects how verbal hatred can have real consequences with consistent escalation, transforming from *rhetoric/speech* to action.

³⁵⁴ Chetty N. Hate speech review in the context of online social networks / N. Chetty, S. Alathur // *Aggression and Violent Behavior*. – 2018. – Vol. 40. – P. 110.

³⁵⁵ Koval E.A. Resentment and resentment as reasons for the use of hate rhetoric in media reality / E.A. Koval // *Information Age*. - 2018. - No. 2-1. (In Russian).

³⁵⁶ Spiridovich E.A. Op. cit. P. 139; Vasilenko E.N. Op. cit. P. 127.

³⁵⁷ Spiridovich E.A. Op. cit. P. 140.

³⁵⁸ Tsesis A. *Destructive Messages: How Hate Speech Paves the Way for Harmful Social Movements* / A. Tsesis. – New York: New York University Press, 2002. – P. 4.

³⁵⁹ Op.cit. (Internet, social media and online hate speech...). P. 4.; Pitruzzella G. Op. cit.

³⁶⁰ Pitruzzella G. *Disinformation and Hate Speech: A European Constitutional Perspective* / G. Pitruzzella, O. Pollicino. – Milano: EGEA Spa - Bocconi University Press, 2020. – P. 31.

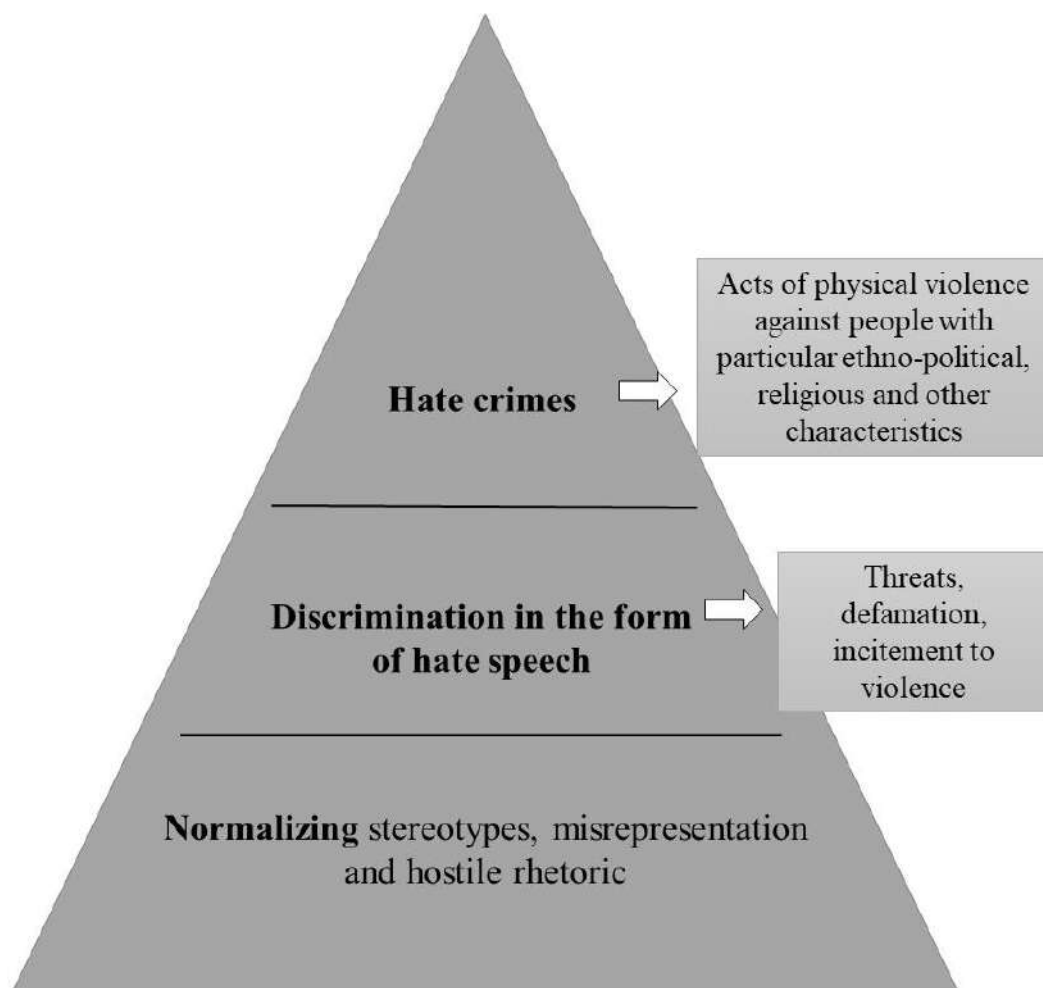


Figure 5 - “Pyramid of hate”

Source: compiled based on data from Pitruzzella G. Op. cit.

In the online space, hate speech takes on the character of a demonstrative public action in the form of a hostile comment or the creation of content that expresses outrage, insult, or direct harassment³⁶¹. Over the past decade, an academic consensus has emerged in hate speech research that social media has become “a fertile ground for heated discussions which frequently result in the use of insulting and offensive language.”³⁶² Although the phenomenon existed earlier in the era of traditional media, the network infrastructure of the Internet has exacerbated the problem. The algorithmic principle of social networks often supports content associated with hate³⁶³ because it provokes a broad

³⁶¹ Shchetinina A.V. Hate And Hatred As Forms Of Enmity In Modern Communication: Experience Of Portraiting Lexema And Concept / A.V. Shchetinina, A.S. Semekhina // Tomsk State Pedagogical University Bulletin. – 2022. – № 1 (219). – P. 38. (In Russian).

³⁶² Thirty years of research into hate speech: topics of interest and their evolution / A. Tontodimamma [et al.] // Scientometrics. – 2021. – Vol. 126. – № 1. – P. 158.

³⁶³ Lim M. Algorithmic Enclaves: Affective Politics and Algorithms in the Neoliberal Social Media Landscape / M. Lim // Affective Politics of Digital Media. – Routledge, 2020.

public response. In addition, by creating echo chambers, new media cultivate a sense of involvement³⁶⁴, making hate speech a successful and actively used communication strategy.

Hate speech varies from academic, legal, or practical perspectives, making it difficult to determine the goals, consequences, and modalities of moderation³⁶⁵. Each actor (platforms, regulators, etc.) distinguishes its characteristic features of hate speech, which creates problems in law enforcement practice.

These contradictions exist at the level of international criminal law³⁶⁶ or the platform rules compliance with national legislation³⁶⁷. In international law, hate speech can be a form of *persecution*, i.e., a crime against humanity, or *incitement*, connecting hate speech to criminal liability³⁶⁸. But at the national level, the interpretations depend on the existing legislative developments, which were partly, and in some cases mainly, formed before the digitalization process. On the other hand, the terms of service and community rules that guide the platforms are often based on arbitrary characteristics of hate speech, reflecting the position of the digital campaign itself.

Despite the pluralism of approaches to the functional definition of hate speech, several features characterize the process of strategic hate dissemination³⁶⁹, including 1) the focus of a *public* message on a particular *group* or an *individual* as a member of a group; 2) *content* expressing hatred; 3) the *intent to harm*; 4) the motivation for *action* and violent reaction.

At the same time, a distinctive feature of hate speech remains a direct relationship with emotions (as a kind of emotional expression)³⁷⁰ and their activation through informational stimuli. In connection with emotionality, the rhetoric of hate shows its social characteristics (*the public nature of the speech act*) and the harmful nature of the

³⁶⁴ Rostova N.N. Philosophy of transmedia: collective monograph / N.N. Rostov. - Moscow: Prospect, 2022. - P. 20-21. (In Russian).

³⁶⁵ Fortuna P. Op. cit. P. 85:4.

³⁶⁶ Fino A. Defining Hate Speech / A. Fino // Journal of International Criminal Justice. – 2020. – Vol. 18. – P. 31-57.

³⁶⁷ Sellars A. Defining Hate Speech / A. Sellars. – Berkman Klein Center Research Publication No. 2016-20. – Boston Univ. School of Law, 2016 – P. 24-31.

³⁶⁸ Fino A. Op. cit. P. 31.

³⁶⁹ Sellars A. Op. cit.

³⁷⁰ Field of Disinformation, Democratic Processes, and Conflict Prevention: A Scan of the Literature. Field of Disinformation, Democratic Processes, and Conflict Prevention / S. Udupa [et al.]. – Social Science Research Council, 2020. – P. 3.

rhetorical action (*possible violent reaction*)³⁷¹ since hate is a strong stimulus and tool for intra-group solidarity. This can be explained by the exploited emotion, i.e., hate, closely related to anger and fear. In the context of emotional regimes and emotives, it is worth considering that such hostile emotions³⁷² have a pronounced social dimension associated with the coexistence of descriptive and performative parameters of hatred. The performative property of emotion provokes other addressees and objects to react to hate speech. While retaining these properties, hate falls under the influence of the emotional regime, which can be subject to external or internal information influence.

Understanding the manipulative nature of emotions means that, under certain circumstances, the social understanding of anger and fear, and thus the derivative of hate, can be changed by a specific influence on the emotional regime, which becomes a threat to the social system and an additional example of information influence at the social level. The subject can incline society towards acceptable hatred by influencing the emotional regime. In the hands of political actors, hate speech (especially in the context of digitalization) becomes a tool for managing perception, leading to serious damage to psychological security at the international level.

In cases of manipulative hate speech, vulnerable social groups become targets for targeted incitement to “mass violations of human rights, crimes against humanity or even genocide.”³⁷³ This phenomenon stems from considerations of maintaining the current discourse (when it implies the need for an object of hate) or, in parallel, circumstances in which the established discourse is an obstacle for political actors, as a result of which the subject aims to transform the established discourse through the hate speech.

Like disinformation, hate speech has its structure of information influence (Figure 6), which functions as an important manipulative strategy of modern digital space. The basis of information influence is the relationship between subject and object, where the object has a dual nature. The primary object is the group of people (with common characteristics of “them”) against whom hatred through hate speech is expressed. The

³⁷¹ Op. cit. (Thirty years of research into hate speech...). P. 158.

³⁷² Ihlebæk K.A. Hostile emotions: An exploratory study of far-right online commenters and their emotional connection to traditional and alternative news media / K.A. Ihlebæk, C.R. Holter // Journalism. – 2021. – Vol. 22. – № 5. – P. 1207-1222.

³⁷³ Fino A. Op. cit. P. 32.

secondary object, similarly to the principle of iteration indicated in the first section of this chapter (Figure 2), acts as a recipient of the primary information influence but, in the case of an emotive connection with the subject, as well as a common position concerning the primary subject, it also becomes a subject within the framework of the replication of hate speech.

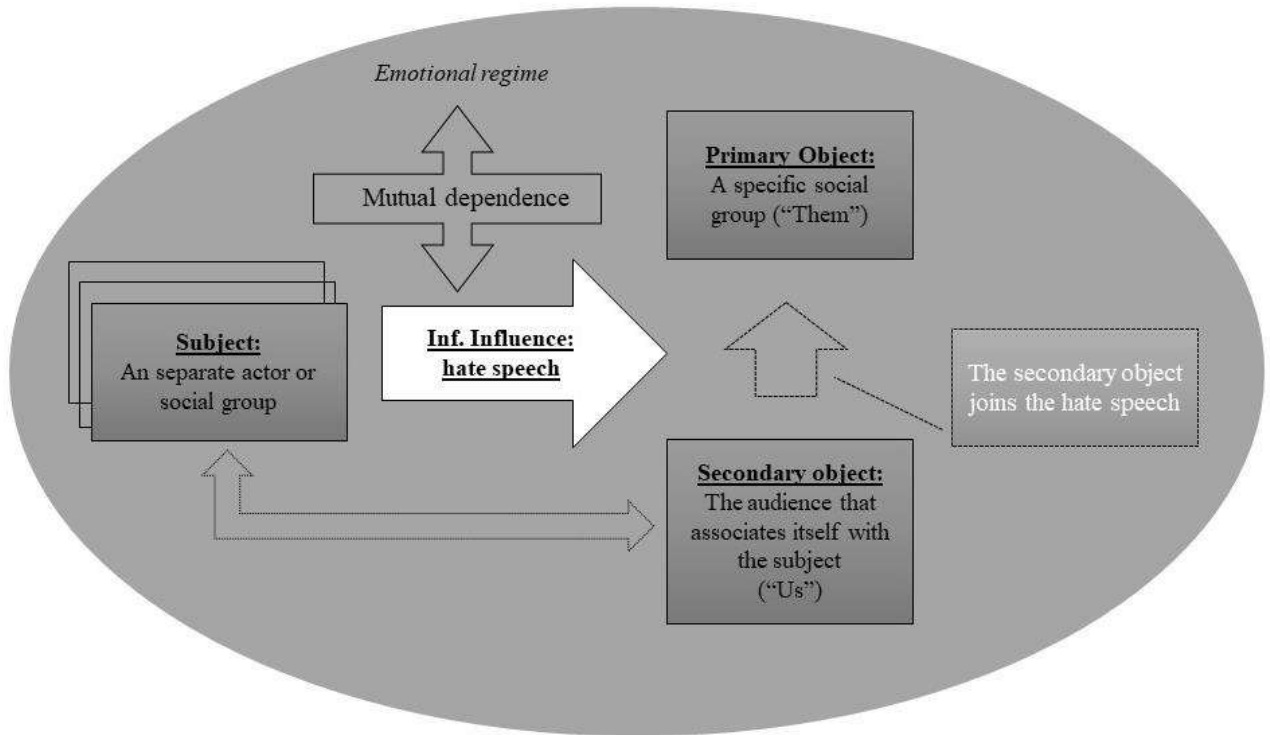


Figure 6 – Subject-object relationships in the spread of hate speech

Source: compiled by the author

In this system, an important circumstance remains the mutual conditioning of the information influence and the emotional regime. All subjects and objects are within the framework of a common emotional regime, which determines the reaction of both primary and secondary objects and the entire society. However, the use of hate speech also has a direct impact on the emotional regime. This aspect also manifests the manipulative property of hate speech, which goes beyond the subject-object relations of a particular individual case of information influence.

Summarizing the provisions on modeling information influence through disinformation and hate speech, it can be observed that in a digital world, “violence has been replaced by manipulation, carried out primarily through the rich resources of

language.”³⁷⁴ The manipulative properties of information influence are simplified by the tools and vulnerabilities of the intermediary in the form of new media. The post-truth era is becoming an accelerator for unverified data dissemination, and the emotionality of the discourse attracts various political actors to exploit the digital space to implement their intentions at both the national and global levels.

At the same time, the structure of information influence shows that the spread of disinformation and hate speech does not always occur under the influence of direct political motivation but also has a situational or subjective character. It poses serious challenges to the analysis and implementation of measures to counter new digital threats, including various forms of prospective manipulative use of automation (bots) and more advanced forms of artificial intelligence (AI), which are becoming one of the most significant factors in the intensification of disinformation and hate speech in the 21st century.

The structural analysis of information influence in the post-truth era has formed a methodological basis to study the interaction of various variables and actors of digital manipulation. The chapter illustrates the relationship between the subject and object of information influence, the role of digital platforms in the international mediation of information messages, and the main tools that make information influence an information threat in world politics.

The sum of all the named variables forms a system in which the subject, the intermediary, and the object remain within a single information environment that shares a multitude of truth discourses and emotional regimes. The environment establishes the primary framework for information exchange. However, in the process of information influence, certain aspects of the information environment may change through

³⁷⁴ Alexeyev A.B. Op. cit. P. 460.

information influence on the object or through the transformation of information during mediation.

The structuring of the object and subject relationships illustrates that each has a methodological framework. The elements of the system have a complex structure of emotional and rational, individual and group preconditions or stimuli. The ultimate manipulation tools are various types of discursive and rhetorical practices, expressed primarily in disinformation and hate speech.

The central mechanism of disinformation is aimed at distortion and false interpretation of data or their falsification to confirm a particular perspective. Hate speech is group sentiments oriented, as it operates with collective emotions, using, for example, aggression and hatred as an effective discourse. In the context of digital reality, both disinformation and hate speech transform into qualitatively new threats to global stability. The systemic representation of information influence formed within the framework of this chapter reflects a set of variables of fundamental importance in terms of the essence of the information manipulation phenomenon and its projection on international reality.

Chapter 3: Information threats in the early 2020s

The growing relevance of the threats identified in the previous chapter is caused not only by their actual status and manipulative nature. The previous analysis illustrated that transformations in information distribution systems have led to a qualitative transformation of the information influence process. At the same time, despite the existing problems related to fakes and hate speech, the relevance of the whole topic increased due to the technologies of the early 2020s, which also have serious potential in information manipulations. One of the central technologies that has illustrated exponential growth over the past few years has been artificial intelligence (AI), in the broad sense of the word, and all its various software and machine applications, including generative neural networks.

AI changes the research agenda of information influence and draws attention to the exceptional dynamics of the information sphere in connection with international relations. The latest stage of AI development, coinciding with the actualization of post-truth politics, forces disinformation researchers to address existing and promising scenarios for adapting the technology to the needs of political manipulation.

The changes in the research agenda also derive from the need to analyze the current experience of countering the latest threats and the attempts to formalize it through more comprehensive management strategies. Although the research community already attempted to study systematically the actions of various national³⁷⁵ and supranational actors³⁷⁶, most of the observations are limited to the consolidation of existing experience, in which theoretical conceptualization takes a secondary position.

³⁷⁵ Sharikov P. US, Europe and Russia Approaches to the Problem of the Information Policy / P. Sharikov, N. Stepanova // *Contemporary Europe*. – 2019. – № 2 (88). – P. 73-83. (In Russian); Fathaigh R.Ó. The perils of legally defining disinformation / R.Ó. Fathaigh, N. Helberger, N. Appelmann // *Internet Policy Review*. – 2021. – Vol. 10. – № 4; Kolotaev Y. Op. cit. (Malicious Information Threats of a Post-Covid World...).

³⁷⁶ Kudrna J. The Possibilities of Combating So-Called Disinformation in the Context of the European Union Legal Framework and of Constitutional Guarantees of Freedom of Expression in the European Union Member States / J. Kudrna // *International Comparative Jurisprudence*. – 2022. – Vol. 8. – № 2. – P. 138-151; Prokopović A.M. The European Approach to Regulating Disinformation / A.M. Prokopović, M. Vujović // *FACTA UNIVERSITATIS - Law and Politics*. – 2020. – Vol. 18. – № 3. – P. 175-183; Sharoichenko N.D. European Union Policy In The Area Of Countering Soft Power Of Russia In The States Of The Eastern Partnership / N.D. Sharoichenko // *Bulletin of Moscow Region State University*. – 2022. – № 2. (In Russian); Kolotaev Yu.Yu. Op. cit. (Evolution Of The EU's Policy...).

This chapter summarizes the experience of major global organizations, various regional actors, and digital platforms within two main paradigms related to protection or adaptation to the new digital reality. Such theoretical frameworks as “resilience” and “multistakeholderism” are also explored, explaining the latest trends in modeling governance activity, including those related to the regulation of information space.

3.1. Artificial intelligence and the future prospects for digital information threats

The problem of information manipulation in international relations since the development of the latest web technologies has initially acquired a technology-oriented character, drawing attention to the new methods of information dissemination. However, since the appearance of fake news and online disinformation issues in the mid-2010s, the technological process has illustrated that the understanding of the issue has evolved. While initially, the main concerns touched online bots and digital platforms with their opaque content delivery methods, nowadays, in addition to these concerns, researchers of online propaganda and manipulation in global politics have, since the late 2010s, begun to shift the focus of their research to AI issues. Although AI was present in a simplified form at the early stages of studying web influence (content delivery algorithms and bots), the development of multimodal data analysis and recognition capabilities with their subsequent self-generation have led to breakthroughs in both the issues of information influence and the prospects for countering it.

AI, as “an artificial complex cybernetic computer-software-hardware system [...] with a cognitive-functional architecture and its own or relevantly available (assigned) computing power of the necessary capacity and speed,”³⁷⁷ is becoming one of the most relevant and advanced information technologies, increasingly spreading into more areas of public life. Unlike other complex computer systems, AI, characterized by the absence of software-defined algorithms, builds upon learning based on the accumulation of

³⁷⁷ Ponkin I.V. Artificial Intelligence from the Point of View of Law / I.V. Ponkin, A.I. Redkina // RUDN Journal of Law. – 2018. – Vol. 22. – № 1. – P. 94-95. (In Russian).

experience³⁷⁸. It makes AI a universal tool for creating complex systems for processing and outputting data.

There are two directions of development and comprehension of AI - artificial general intelligence and weak or narrow AI. Weak AI is “actively used in many industries, from browsers [...] to driving cars and airplanes.”³⁷⁹ General AI, in its truest sense, has not yet emerged and represents a hypothetical model of AI that should have cognitive abilities equal to or even superior to humans. For this reason, information manipulation researchers consider weak and narrow forms of AI.

To date, “the application of AI [...] raises important issues of information security in diplomacy, the protection of personal data in social networks, and the mitigation of the effects of algorithmic bias.”³⁸⁰ This trends link to the broader phenomenon of the datafication³⁸¹ of international relations or the widespread use of big data analytics for political purposes.

As a result, a new information threat associated with the use of AI for political manipulation, i.e., threat of the malicious use of artificial intelligence (MUAI)³⁸², is gradually emerging within the framework of world politics. Such use of AI implies causing social damage at the level of public perception, infrastructure, or psychological security. Within this problem field, several levels at which AI can become a threat to society are distinguished. Prof. E.N. Pashentsev distinguishes three threat levels to psychological security created by MUAI³⁸³. Their manifestation and the degree of damage depend on the specific form of AI.

At the first level, as a result of any incidents, society develops a negative reaction to AI, which causes a decrease in its integration into social development and social processes. The second level is related to direct MUAI, aimed at destroying or damaging

³⁷⁸ Ponkin I.V. Op. cit. P.94.

³⁷⁹ Korobkov A.D. The Impact of Artificial Intelligence Technologies on International Relations / A.D. Korobkov // MGIMO Review of International Relations. – 2021 – P. 2. (In Russian).

³⁸⁰ Grishanina T.A. Artificial intelligence in international relations: role and research dimensions / T.A. Grishanina // RSUH/RGGU Bulletin Series "Political Science. History. International Relations". – 2021. – № 4. – P. 17. (In Russian).

³⁸¹ Tsvetkova N.A. Op. cit.

³⁸² Pashentsev E. Op. cit. P. 33; Experts on the Malicious Use of Artificial Intelligence: Challenges for Political Stability and International Psychological Security / ed. E. Pashentsev. – Moscow: International Center for Social and Political Studies and Consulting, 2020; Kolotaev Y. Op. cit. (Methods for Countering ...).

³⁸³ Bazarkina D. Malicious Use of Artificial Intelligence / D. Bazarkina, E. Pashentsev // Russia in Global Affairs. – 2020. – Vol. 18. – № 4. – P. 154-177.

infrastructure, causing social harm, causing financial damage, etc. Negative emotions, anxiety, and panic moods inevitably appear. However, it is only the third level of threat that involves distorting and controlling perceptions or causing psychological damage. It is the third level where the use of AI in psychological confrontation or even information warfare occurs. All three levels help to understand how a social system can respond to different types of MUAI.

Of all the three levels mentioned above, the third, related to information threats, reflects the most promising area for exploiting the negative aspects of AI. As mentioned above, some aspects affect the platform activity since content distribution and targeting algorithms are already built on vulnerable systems, representing weak AI exploited in content generation. However, there are also those areas of AI that are only at the initial stage of integration into reality but still may become significant tools in the hands of information influence actors.

The context of the post-truth era and the increasing emotionality of public discourse make us pay special attention to the field of emotional AI³⁸⁴ and automated sentiment analysis³⁸⁵, which are more broadly related to affective computing³⁸⁶. Emotional AI and affective computing are emerging interdisciplinary fields that explore how AI and other computing technologies can make it easier to understand human emotions. It also addresses human-computer interaction issues, e.g., the design of systems that apply emotion recognition to augment their functionality.

This range of technologies indicates either the ability to recognize human emotions for the subsequent optimization of information systems or to simulate emotions in such information systems. Emotion in such a context are functionally interpreted as “linguistic or non-linguistic expression of feeling, [...] passions, sentiments and drives, in words,

³⁸⁴ McStay A. Emotional AI: The Rise of Empathic Media. *Emotional AI* / A. McStay. – First edition. – Los Angeles: SAGE Publications Ltd, 2018. – 248 p.; Stark L. The Ethics of Emotion in Artificial Intelligence Systems / L. Stark, J. Hoey // *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency: FAccT '21*. – New York, NY, USA: Association for Computing Machinery, 2021. – P. 782-793; Emotion AI-Driven Sentiment Analysis: A Survey, Future Research Directions, and Open Issues / P. Chakriswaran [et al.] // *Applied Sciences*. – 2019. – Vol. 9. – № 24. – P. 5462.

³⁸⁵ Yadollahi A. Current State of Text Sentiment Analysis from Opinion to Emotion Mining / A. Yadollahi, A.G. Shahraki, O.R. Zaiane // *ACM Computing Surveys*. – 2017. – Vol. 50. – № 2. – P. 25:1-25:33; A survey of sentiment analysis in social media / L. Yue [et al.] // *Knowledge and Information Systems*. – 2019. – Vol. 60. – № 2. – P. 617-663.

³⁸⁶ Picard R.W. *Affective Computing* / R.W. Picard. – Cambridge, MA, USA: MIT Press, 1997. – 306 p.; Deep learning for affective computing: Text-based emotion recognition in decision support / B. Kratzwald [et al.] // *Decision Support Systems*. – 2018. – Vol. 115. – P. 24-35.

gestures and other social practices.”³⁸⁷ Emotion recognition through intelligent systems occurs both at the level of a text and a non-textual or multimodal source³⁸⁸. It means that it is possible to interpret human sentiments within the decision-making process of the AI system itself. The simulation function complements this possibility by imitating quasi-human reactions to incoming emotional stimuli from the user in a textual and multimodal format.

Despite the great potential of emotional AI in general, the development of technology is associated with risks of its use within the framework of information influence³⁸⁹ or in the course of full-fledged political campaigns to control perception³⁹⁰. The application of emotional computing for political purposes turns sentiment analysis and emotional AI into “digital weapons.” Both tools, for example, allow the extraction and analysis of human emotions for precise empathic targeting of information influence. In such a context, human emotions become subject to various forms of exploitation through AI algorithms in the manipulation process. These include both empathic simulation and the recognition of feelings through specialized markers. The promising effectiveness of such methods derive from the role of emotions, impacting significantly human performance and cognitive processes. As emotions remain one of the most important aspects of perception, the possibility of instrumentalizing or altering them creates severe societal and individual security implications.

There are multiple potential scenarios where the malicious use of emotional AI could become an information weapon for manipulative strategies that use precise empathic targeting. These include emotionally optimized disinformation³⁹¹, large-scale

³⁸⁷ Early Modern Emotions: An Introduction. *Early Modern Emotions* / Ed. S. Broomhall. – London: Routledge, 2016. – P. 11.

³⁸⁸ Chakraborty K. A Survey of Sentiment Analysis from Social Media Data / K. Chakraborty, S. Bhattacharyya, R. Bag // *IEEE Transactions on Computational Social Systems*. – 2020. – Vol. 7. – № 2. – P. 450-464.

³⁸⁹ Pantsev K.A. Op. cit.; Roumate F. Malicious Use of Artificial Intelligence, New Challenges for Diplomacy and International Psychological Security / F. Roumate // *Artificial Intelligence and Digital Diplomacy*. – 2021. – P. 97-113; *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. The Malicious Use of Artificial Intelligence* / M. Brundage [et al.] arXiv: 1802.07228. – 2018.

³⁹⁰ Mengü M.M. Digital Perception Management / M.M. Mengü, E. Türk // *Digital Seige*. – Istanbul, Turkey: Istanbul University Press, 2021. – P. 327; Derman G.S. Perception management in the media / G.S. Derman // *International Journal of Social and Economic Sciences*. – 2021. – Vol. 11. – № 1. – P. 64-78; Kolotaev Y. Op. cit. (*Sentiment Analysis...*); Bakir V. Psychological Operations in Digital Political Campaigns: Assessing Cambridge Analytica’s Psychographic Profiling and Targeting / V. Bakir // *Frontiers in Communication*. – 2020. – Vol. 5.

³⁹¹ AI and the Future of Disinformation Campaigns / Center for Security and Emerging Technology [Electronic resource]. – URL: <https://cset.georgetown.edu/publication/ai-and-the-future-of-disinformation-campaigns/> (accessed: 12.02.2023).

propaganda campaigns through social engineering³⁹², hate speech incitement through social bots³⁹³, and information espionage based on emotional profiling³⁹⁴ or spear phishing. Moreover, all these scenarios aim at one of two information system vulnerabilities³⁹⁵: algorithm vulnerability (internal information system vulnerability) and user vulnerability (external to the intermediary) vulnerability (Figure 7).

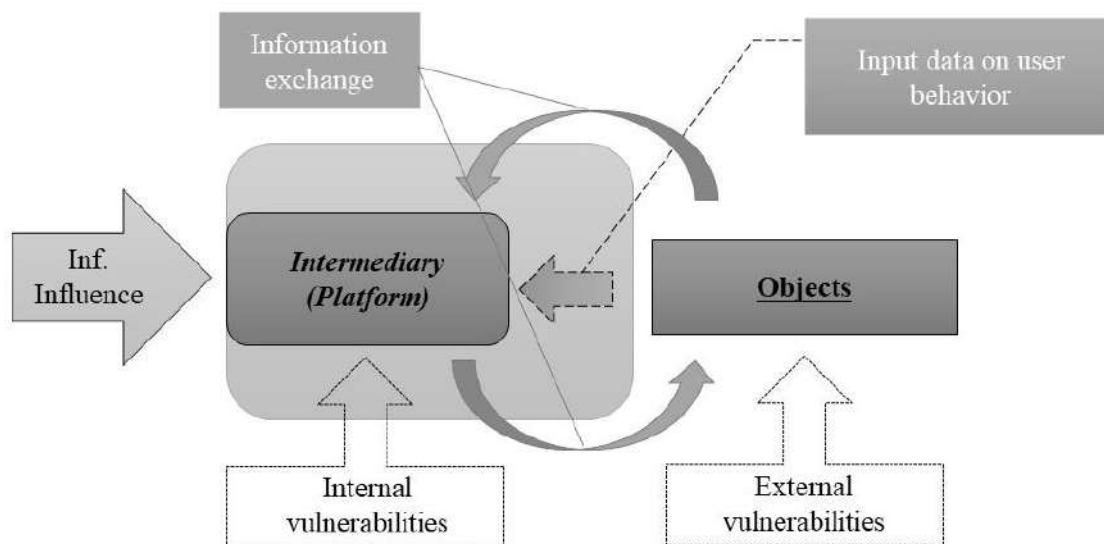


Figure 7 – Forms of information system vulnerability

Source: compiled by the author

The first form of vulnerability involves, for example, a distortion of the data processing mechanism for modifying results. In such a case, the influence subject manipulates the input data to change the results of the sentiment classifier (if talking about sentiment analysis). Such an information attack aims indirectly at the influence object. It does not directly threaten its information field but distorts the AI-based tools used by the object during decision-making. Consequently, it can lead to situations in which incorrect decisions are made³⁹⁶ or distorted results are produced. The targets of such attacks are

³⁹² Djugumanov I. Practical Applications of Advanced Technologies for Enhancing Security and Defense Capabilities: Perspectives and Challenges for the Western Balkans / I. Djugumanov, M. Hadji-Janev. – NATO Science for Peace and Security Series-E: Human and Societal Dynamics. – IOS Press, 2022. – 134 p.

³⁹³ Social Bots and the Spread of Disinformation in Social Media: The Challenges of Artificial Intelligence / N. Hajli [et al.] // British Journal of Management. – 2022. – Vol. 33. – № 3. – P. 1238-1253.

³⁹⁴ Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media / B. Zarouali [et al.] // Communication Research. – 2022. – Vol. 49. – № 8. – P. 1066-1091.

³⁹⁵ Kolotaev Y. Op. cit. (Sentiment Analysis...).

³⁹⁶ Word-level Textual Adversarial Attacking as Combinatorial Optimization / Y. Zang [et al.] // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics / arXiv:1910.12196 [cs]. – 2020. – P. 6066-6080.

either machine learning algorithms (at the level of emotional AI)³⁹⁷ or lexical sentiment analysis classifiers (at the level of sentiment analysis)³⁹⁸, leading to “data poisoning” or interference in language processing algorithms.

The second user-oriented form involves the direct use of emotional AI against humans in contextually and emotionally optimized information attacks. Such attacks combine phishing, social engineering³⁹⁹, psychological profiling⁴⁰⁰, and social bot attacks⁴⁰¹. This form of information influence is aimed directly at the object and uses data about its emotional state, response, or perception of the current issue or topic. Such attacks involve access to personal or sensitive data of the user(s) that optimize the attack results. This targeted use of emotional AI and, especially, automated sentiment analysis leads to the highly precise and targeted spread of hate speech, fakes, and other forms of disinformation. The most explicit example in this regard is empathically optimized automated fake news⁴⁰². It mostly remains only a promising form of information influence. Still, it can adjust information to the target audience by analyzing the previous emotional response, allowing it to specify the scale and results of information influence. This scenario is especially relevant for isolated communities (existing in echo chambers), which have a predetermined range of opinions and forms of emotional response.

While this form of information influence remains future-oriented, its examples of perception management campaigns (with attempts to use data-based targeting) confirm its actuality. The most prominent example was the Cambridge Analytica scandal, involved in the UK’s Brexit referendum⁴⁰³. Further modifying similar campaigns through emotional AI is a matter of evolving the technology itself, especially in the face of growing digital conflicts and information wars. The early 2020s show that manipulative

³⁹⁷ Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning. *Manipulating Machine Learning* / M. Jagielski [et al.] arXiv:1804.00308 [cs]. – arXiv, 2021.

³⁹⁸ Adversarial attacks on a lexical sentiment analysis classifier / G.A. de Oliveira [et al.] // *Computer Communications*. – 2021. – Vol. 174. – P. 154-171.

³⁹⁹ The role of Social Engineering in evolution of attacks / E. Frumento [et al.]. – DOGANA Consortium, 2016.

⁴⁰⁰ Jahankhani H. *Cybercrime Profiling and Trend Analysis* / H. Jahankhani, A. Al-Nemrat // *Intelligence Management: Knowledge Driven Frameworks for Combating Terrorism and Organized Crime* / eds. B. Akhgar, S. Yates. – London: Springer, 2011. – P. 181-197.

⁴⁰¹ Social Robot Interactions for Social Engineering: Opportunities and Open Issues / A.F. Abate [et al.] // *2020 IEEE Intl Conf (DASC/PiCom/CBDCCom/CyberSciTech)*. – 2020. – P. 539-547.

⁴⁰² Bakir V. Op. cit. (Fake News...).

⁴⁰³ Bakir V. Op. cit. (Psychological Operations...).

information campaigns are becoming more common and accessible⁴⁰⁴, and, therefore, emotional automated confrontation tools require special attention, as the growing availability of AI has severe implications for the future.

The increasing availability of emotion recognition systems creates other multilayered psychological attack scenarios. For example, an AI-assisted perception management operation can be modeled through several sequential steps (Table 3). A subject plans an attack on a particular social group. In the first step, he applies sentiment analysis to gain a precise *understanding of the basic emotional regime of the community*. It facilitates further planning of the entire operation and enables the subject to target the community's emotional triggers. In this way, the actor creates the necessary *tools to infiltrate various communities* through social media to spread disinformation or introduce elements of hate speech into the target social group. The tools can be advanced social/emotional bots that spread hate messages by adapting them to the discourse of the social group.

In the second stage, the audience's *reaction analysis* to the disseminated data is critical. It helps to amplify the original message and opens up opportunities to continue the information influence on an iterative basis. *Each iteration* will include new data on emotional reactions, making the overall attack more targeted. If the subject has sufficient computational ability, the next step may be the creation of *synthetic, emotionally adapted fakes* already mentioned. In this phase, the attack will rely more on the AI system itself, reducing the role of the human agent and creating a continuous *system of dissemination and subsequent analysis*. Once the community accepts the main message, and thus the idea of information influence, the actor can expand the bot network or change the original target if further manipulation is needed.

Table 3 – Multi-level scenario of psychological attacks using AI

Stage	1) Planning
1.1	Understanding the basic emotional regime of the community

⁴⁰⁴ Kolotaev Y. Op. cit. (Hate Speech in Perception Management Campaigns); Op. cit. (AI and the Future of Disinformation Campaigns).

Table 3 continued

Stage	1) Planning
1.2	Identifying the community's emotional triggers
1.3	Creating the necessary tools to reach the community
	2) Information dissemination and analysis
2.1	Application of the tools developed for information dissemination
2.2	Analyzing audience reaction to disseminated data
2.3	Reinforcing the original message to continue information influence
	3) New iteration of information implementation and analysis
3.1	New data on emotional reactions for a more targeted attack
3.2	Creating synthetic, emotionally adapted fakes
3.3	Establishment of a continuous system of dissemination and subsequent analysis
	4) After achieving the goal
4.1	Expanding the bot network to increase the sphere of influence or selecting a new target

Source: compiled by the author

Such an abstract modeled example only predicts the role emotional AI may eventually play in perception management. Still, the development of multiple new models of potential AI attacks developed by the scientific community illustrates the relevance and urgency of the threat. In particular, these include various types of synthetic audiovisual disinformation and hate speech used to discredit or deceive⁴⁰⁵. In a broader sense, they are called deepfakes and refer to fake video images generated by machine learning that simulate the overlay of one image on another⁴⁰⁶. Due to its visual features, this form of AI application has received considerable media attention, which has led to

⁴⁰⁵ Pawelec M. Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions / M. Pawelec // Digital Society. – 2022. – Vol. 1. – № 2. – P. 19.

⁴⁰⁶ Brown N.I. Deepfakes and the Weaponization of Disinformation / N.I. Brown // Virginia Journal of Law & Technology. – 2020. – Vol. 23. – P. 1; Vaccari C. Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News / C. Vaccari, A. Chadwick // Social Media + Society. – 2020. – Vol. 6. – № 1.

deepfakes being the most actively problematized form of AI compared to other forms of AI in the context of information manipulation.

Deepfakes raise a serious problem, first of all, due to the supposedly credible information received. On the one hand, deepfakes become a practical tool for political deception. On the other hand, in case of disagreement with the information source, a political actor can refer to the falsity of images or data, making the problem of deepfakes a rhetorical instrument, similar to what happened with “fake news” and “fakes.”

It is worth noting, however, that deepfakes, like other forms of AI, do not pose an absolute threat to the information space. They only become such as an instrument of information influence. This circumstance does not negate the great potential of AI in its various forms. The aforementioned automated sentiment analysis and emotional AI are also applicable in countering information threats⁴⁰⁷. Given the nature of AI to act through a trained model, the automated system itself may aim to identify untrustworthy content through various distinctive direct and indirect indicators⁴⁰⁸. For this reason, AI should not be seen as a fundamentally new information influence threat but only as a way to strengthen or counteract it. However, the danger of excessive optimism about AI’s ability to recognize unreliable information is often overlooked. The main limitation of AI remains both the representation of AI training samples and the applicability of most human cognitive biases to AI since the algorithm formation highly depends on the vision of algorithm creators. Thus, the gap between general AI and weak AI ensures that AI practices are imperfect in both disseminating and countering misinformation.

Important to mention that the development of response mechanisms to counter threats is in a catch-up state of reactive threat response. The threat itself is proactive and prior. In addition, the concept of “weak AI” means that it is limited relative to general AI. In narrow spheres, it can deliver high performance that exceeds the results of actions without AI-based tools. The spheres of information influence and psychological manipulation progressively confirm it as well.

⁴⁰⁷ Op. cit. (Sentiment Analysis for Fake News Detection). P. 1348; Fake News Detection Using Sentiment Analysis / B. Bhutani [et al.] // 2019 Twelfth International Conference on Contemporary Computing (IC3). – 2019. – P. 1-5.

⁴⁰⁸ Hate speech operationalization: a preliminary examination of hate speech indicators and their structure / J. Papcunová [et al.] // Complex & Intelligent Systems. – 2021.

To sum up, the new digital threats that contribute to the spread of disinformation and hate speech, we should again emphasize the applicability of AI in information manipulation through content algorithms, the creation of synthetic data, and the emotional state analysis of users. The AI systems' integration into new media is of particular concern, as it transforms the system for providing new information and its substance. In particular, AI enables the targeted selection of relevant content and the optimization of data for a specific audience, which, in the context of information influence, becomes an amplifying factor of the threat itself.

Thus, when studying information influence, it is important to understand the influence tools themselves, e.g., hate speech or disinformation, and to analyze future-oriented ways to improve these tools in modern conditions. The evolution of information dissemination tools plays a crucial role in the appearance of the post-truth era in the 21st century. For this reason, understanding the negative consequences of such conditions is possible only if one accepts the importance of both the initial assumptions about the problems of truth and emotionality in society and the characteristics of modern technological development, which introduces new, revolutionary tools of social communication through AI and its connection with social networks.

3.2. Actors and models of countering information threats in international relations

The transformation of information influence into new types of information threats, which takes place in the digital environment, poses the task to several actors to find protection means, methods, and mechanisms. World practice illustrates different approaches depending on the level of counteraction and the relations between the system and the threat. The spatial and political understanding of social and international relations determines the existence of *global*, *regional*, *national*, and *local* levels of world politics⁴⁰⁹. This division corresponds to the nature of the institutions of the international

⁴⁰⁹ Tumanov A.D. The Analysis of Factors of Formation of the Modern System of Political Governance: Global, Regional and Local Levels / A.D. Tumanov // Administrative consulting. – 2021. – Vol. 151. – № 7. – P. 107-115. (In Russian); Gukasov A.V. Systemic and subsystemic levels of foreign policy analysis of modern states / A.V. Gukasov // A new word in science and practice: hypotheses and approbation of research results. - 2014. - No. 12. - P. 83-88. (In Russian).

system but contradicts in a certain way with the essence of the digital space characterized by the decentralization of connections (in the horizontal and vertical understanding of spatiality).

However, from a formal point of view, the reaction of political actors to information threats today corresponds to the spatial levels of the international system, which are of great importance for understanding the functional characteristics of various actors in the regulation of digital space and also indicate the potential for fragmentation of the global network. Each level has its actors, including the state and its institutions, international and national organizations, transnational corporations, the media, and civil society. The information challenges have become, for most actors, a socio-political threat that requires discursive, regulatory, educational, or technical interventions. The response form is mostly dependent on the actor and its role within the international system.

The regional distribution of actors is particularly important since they may interpret contemporary digital threats differently due to the political situation in the region of their territorial location. The pluralism of practices and actors creates the need to establish regional or actor-based research frameworks to reflect countermeasures substantively. Attempts to combine actors from different regions with a systemic and spatial representation of global politics can lead to a distorted view of the involved degree of various actors. For this reason, the presented actors and counteraction models are regionally specific and predominantly centered on the Eurasian space.

At the same time, despite the regional specifics, the *global level*, defined in central institutions of the international system, plays a distinctive role in structuring the practices of countering information threats. The first manifestations of information threat regulation at the global level appeared in the mid-2010s. The main actor at this level was the UN, which, following the general trend, gradually began to express its position on combating disinformation and digital hate speech. In 2017, UN Secretary-General A. Guterres stated the need to protect free media to uphold the “right to truth,” as “this is

crucial to counter prevailing misinformation.”⁴¹⁰ In 2019, the Secretary-General further noted that “new technologies are sometimes used to mislead public opinion,” while the spread of misinformation “poses a threat to people’s lives, health security and to public health systems across the world,”⁴¹¹ as demonstrated by the example of misinformation about vaccinations and the Ebola epidemic.

It is important to note that such statements were made one year before the worldwide COVID-19 pandemic in 2020, starting an infodemic associated with misinformation about a new coronavirus caused by what the World Health Organization (WHO) interprets as “too much information including false or misleading information in digital and physical environments during a disease outbreak.”⁴¹² In this context, WHO recommends four types of activities to address the infodemic: raising awareness of community concerns and issues, promoting understanding of risks and advice from health experts, building *resilience* to misinformation, and engaging and empowering communities to take positive actions.

A more detailed vision of the UN’s fight against disinformation appeared in the 2021 General Assembly Resolution on Countering Disinformation⁴¹³ and the 2022 Report of the Secretary-General⁴¹⁴. The resolution notes the need for *civil society and media* efforts to raise awareness of disinformation; calls on states to counter all forms of disinformation through education, building *resilience* to disinformation and raising awareness; recalls the obligations of *companies and platforms* to respect human rights; and calls on all *stakeholders* to promote digital literacy as a way to empower all people. The Secretary-General’s report reinforces these statements with recommendations to governments and technology companies calling for greater transparency, active engagement of diverse communities from the global to the local level, respect for human

⁴¹⁰ In “post truth” era, leaders must defend objective, independent media, UN says on Press Freedom Day / UN News [Electronic resource]. – URL: <https://news.un.org/en/story/2017/05/556472-post-truth-era-leaders-must-defend-objective-independent-media-un-says-press> (accessed: 04.12.2022).

⁴¹¹ Inaccurate Information Makes Effective Responses to Complex Health Crises More Difficult, Secretary-General Tells Event on Countering ‘Fake News’ / UN Press [Electronic resource]. – URL: <https://press.un.org/en/2019/sgsm19574.doc.htm> (accessed: 04.12.2022).

⁴¹² Infodemic / WHO [Electronic resource]. – URL: <https://www.who.int/health-topics/infodemic> (accessed: 04.12.2022).

⁴¹³ United Nations. Op. cit.

⁴¹⁴ Secretary-General UN. Countering disinformation and promotion and protection of human rights and fundamental freedoms / UN. Secretary-General // Publisher: UN. – 2022.

rights, increased public education, and openness to study and public scrutiny of both government policies and platform practices.

In broad terms, the UN framework for action focuses on multidimensional responses consistent with international law and the interests of diverse stakeholders. Special attention is given to education, as evidenced by earlier recommendations made by the United Nations Educational, Scientific and Cultural Organization (UNESCO)⁴¹⁵. From a systemic perspective, this emphasizes the protection of the information influence object and cooperation with the intermediary while ignoring the subject.

Similar recommendations and trends appear in the 2019 UN Strategy and Action Plan to combat hate speech⁴¹⁶. It is important, that hate speech, due to its greater connection with international law and human rights, received a strategic document addressing measures to counter the threat earlier than disinformation. In this regard, the 2021 Resolution refers to the 2019 Strategy, noting that “hate speech and disinformation are distinct phenomena which may overlap in some cases,”⁴¹⁷ as seen during the infodemic⁴¹⁸. In terms of content, the Hate Speech Strategy also draws attention to the importance of interaction between stakeholders, teachers, and researchers while separately highlighting support for victims of hate speech. The provisions of the Strategy were further developed in the General Assembly Resolution “Promoting interreligious and intercultural dialogue and tolerance in countering hate speech,”⁴¹⁹ adopted in 2021, which decided to proclaim 18 June as the International Day for Countering Hate Speech. These UN documents and, particularly, the Resolution have received a positive response and support from Russia.

In addition to the UN, other global and *regional international organizations* and alliances existing in the European and Eurasian space are also taking initiatives in

⁴¹⁵ Journalism, fake news & disinformation: handbook for journalism education and training (In Russian) / UNESCO Digital Library [Electronic resource]. – URL: <https://unesdoc.unesco.org/ark:/48223/pf0000371785> (accessed: 04.12.2022).

⁴¹⁶ Op. cit. (The UN Strategy and Plan of Action).

⁴¹⁷ United Nations. Op. cit. P. 2.

⁴¹⁸ Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation / WHO [Electronic resource]. – URL: <https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation> (accessed: 04.12.2022).

⁴¹⁹ Resolution "Promoting interreligious and intercultural dialogue and tolerance in countering hate speech" A/RES/75/309 / UN General Assembly. – 2021.

connection with the growth of disinformation and hate speech, engaging in digital regulation. At the same time, they differ by various degrees of elaboration of initiatives, and some of them are only at the initial stage of articulating the need for a joint supranational response to threats. In particular, within the framework of the CSTO, the first prerequisites for joint action in the sphere of information threats emerged in the public sphere only in mid-2022⁴²⁰, being a preliminary conceptualization of a common approach in information wars, as well as the need for a joint strategic narrative.

Another significant interregional actor - the CIS - similarly did not institutionalize its approach to disinformation and online manipulation problems yet. The only documents and bodies that serve as a basis for potential actions in this area are the 2013 Agreement on Cooperation of CIS Member States in the Field of Information Security⁴²¹ and separate actions of the CIS Anti-Terrorist Center⁴²². At the same time, the Agreement remains only a framework document summarizing the essential desire for cooperation in the field of information threats, while the activities of the Anti-Terrorist Center are limited to analytical activity⁴²³. Thus, despite the pluralism of regional initiatives of international institutions, we can observe different degrees of conceptualization and elaboration of measures and actions in combating disinformation, leading to divergent attitudes toward digital threats.

In parallel with, and some cases ahead of, international institutions, *national actors* have joined the fight against digital threats. In recent years, various countries have taken legislative initiatives to prohibit or increase liability for fake and hateful content, in some cases, by creating new norms, and in others, by applying existing legal practices to the digital space.

⁴²⁰ On May 16, a meeting of the leaders of the CSTO member states was held in Moscow, dedicated to the 30th anniversary of the signing of the Collective Security Treaty and the 20th anniversary of the creation of the CSTO. The Statement of the members of the Collective Security Council was adopted (In Russian) [Electronic resource]. – URL: https://odkb-csto.org/news/news_odkb/v-moskve-16-maya-proydet-vstrecha-glav-gosudarstv-chlenov-odkb-posvyashchennaya-30-letiyu-podpisaniya/#loaded (accessed : 05.03.2023).

⁴²¹ On the signing of the Agreement on Cooperation of the Member States of the Commonwealth of Independent States in the field of information security / Ministry of Digital Development, Communications and Mass Media of the Russian Federation [Electronic resource]. – URL: https://digital.gov.ru/ru/documents/3729/?utm_referrer=https%3a%2f%2fwww.google.com%2f (accessed: 09.09.2023).

⁴²² Information and Analytical Activity / Anti-Terrorist Center of the CIS Member States [Electronic resource]. - URL: <https://www.cisatc.org/1289/134> (accessed: 09.09.2023).

⁴²³ “Deep Fakes”: Concept and Assessment of Potential Impact on National Security. Analytical note. - Moscow: ATC CIS, 2019.

The Russian experience illustrates lawmaking practices related to the criminalization of various forms of misleading information within the framework of established legislative norms. The initial initiative was the 2019 law on the dissemination of malicious information⁴²⁴, representing essentially a set of amendments to the Federal Law on Information, Information Technologies, and Information Protection and the Code of the Russian Federation on Administrative Offenses. The amendments aim to establish administrative responsibility for the deliberate dissemination of inaccurate information of public importance under the guise of reliable messages, which threatens harm. The interpretation of the law had a wide field of application, but it was supplemented and reinforced as new threats emerged.

Thus, in the spring of 2020, in the context of an infodemic in Russia, amendments were made to the Criminal Code of the Russian Federation in connection with the threat of deliberate dissemination of false information about circumstances dangerous to the life and health of the population⁴²⁵. The purpose of these amendments was to reduce panic from disinformation by increasing the level of responsibility of citizens for the published information about the COVID-19 pandemic. Further practice, which followed in the first half of the 2020s accompanied by growing global confrontation, showed an increasing trend to criminalize the spread of fakes within the framework of criminal law.

The national specificity of other countries also illustrates the diversity of scope, orientation, and the degree of obligation of laws related to countering information influence⁴²⁶. But in addition to laws, national regulatory practices around the world include a range of measures, including “information correction, content removal or blocking, and criminal sanctions.”⁴²⁷ In connection with varying practices, there are discussions about the excessiveness or, vice versa, the inadequacy of the measures

⁴²⁴ Federal Law of 18.03.2019 No. 31-FZ on Amending Article 15-3 of the Federal Law on Information, Information Technologies and Information Protection [Electronic resource]. - URL: <http://publication.pravo.gov.ru/Document/View/0001201903180031> (accessed: 15.09.2023); Federal Law of 18.03.2019 № 27-FZ on Amending the Code of the Russian Federation on Administrative Offenses. [Electronic resource]. - URL: <http://publication.pravo.gov.ru/Document/View/0001201903180021?index=1> (accessed: 15.09.2023).

⁴²⁵ Federal Law of 01.04.2020 № 100-FZ on Amendments to the Criminal Code of the Russian Federation and Articles 31 and 151 of the Code of Criminal Procedure of the Russian Federation. [Electronic resource]. - URL: <http://publication.pravo.gov.ru/Document/View/0001202004010073> (accessed: 15.09.2023).

⁴²⁶ Member State Actions / EU DisinfoLab [Electronic resource]. – URL: <https://www.disinfo.eu/resources/member-state-actions/> (accessed: 20.12.2022).

⁴²⁷ Helm R.K. Op. cit.

taken⁴²⁸. In their extreme forms, they either do not bring results or can cause additional harm to the public.

In particular, concerning the restrictive nature of countermeasures, researchers focus on the threat to freedom of speech and expression⁴²⁹, which derives not as much from governmental actions as those of platforms. Moreover, a certain degree of restriction of freedom of expression is an inevitable feature of the legal fight against disinformation since it is necessary to prevent the spread of false information legally⁴³⁰.

In addition to the opposition to the degree of coercion and rigidity of the applied national practices, another issue is the definition of the direct responsibility of the parties to the information exchange. In other words, the dilemma arises as to who bears the responsibility for the spread of hostility and disinformation: the platforms (as intermediaries) or the users themselves (i.e., the subject of influence). This issue is presented differently in different jurisdictions, depending on the orientation of the law. European practice tends to focus more on the responsibility of platforms as intermediaries in the information space. However, there are alternative views⁴³¹ that call for the criminalization of online disinformation for ordinary users.

The *platforms* themselves and the companies behind them are also independently involved in the process of responding to digital threats. The basis of their response is the accepted terms for using the service, community rules, user agreements, and active measures to intervene in the distribution of content in the form of algorithmic or human moderation. The normative practice associated with the accepted rules for operating the platform and interacting with users is the public side of the actions taken by online services.

For example, most major global social networks openly publish their stance on fakes and hate speech, demonstrating their content management principles. For example, the video hosting YouTube specifies in the community rules a ban on discriminatory

⁴²⁸ Lohani A. Op. cit.

⁴²⁹ Spiridovich E.A. Op. cit. P. 140; Vese D. Governing Fake News: The Regulation of Social Media and the Right to Freedom of Expression in the Era of Emergency / D. Vese // European Journal of Risk Regulation (EJRR). – 2022. – Vol. 13. – P. 477.

⁴³⁰ Helm R.K. Op. cit. P. 303.

⁴³¹ France debates making fake news a crime / Acepresa [Electronic resource]. – URL: <https://www.acepresa.com/english/france-debates-making-fake-news-a-crime/> (accessed: 20.12.2022).

material, false medical information, and other destructive content⁴³². At the same time, the video hosting has a separate anti-hate rhetoric policy that includes rules and principles that, if violated, will lead to content blocking⁴³³.

Other global platforms and companies adhere to similar policies and principles, although they may differ in the understanding and interpretation of the content in question and implementation practices. For example, the Telegram messenger's basic restrictions relate to violent content, while less violent forms of hate speech or false information are not included in its user agreements⁴³⁴. The social network TikTok, on the other hand, includes in its community rules⁴³⁵ a ban on the publication or distribution of disinformation and incitement to hatred, stating that disinformation must cause "significant harm" (specifying what is meant by this) and clarifying the areas in which disinformation is not allowed (medicine, elections, civil disputes, etc.).

In the second half of the 2010s and early 2020s, along with open platform rules and official platform policies, new collaborative approaches to the involvement and cooperation of platforms in the fight against disinformation emerged in international practice. Such measures include the appearance in 2018 of the first internationally recognized Code of Practice on Combating Disinformation, which functions in the European space⁴³⁶.

The Code, unlike other actions, is an independent initiative of online platforms. For the first time, representatives of very large social networks and online platforms (Google, etc.) agreed voluntarily to develop common self-regulatory standards to combat disinformation. The Code established 21 commitments in areas ranging from transparency in political advertising and fighting fake accounts to user involvement in moderation and demonetizing disinformation spreaders⁴³⁷. The Code design consolidates the practical aspects of cross-sectoral collaboration (multistakeholderism) since it

⁴³² Op. cit. (YouTube Community Guidelines...).

⁴³³ Op. cit. (YouTube. Hate speech policy).

⁴³⁴ Op. cit. (Terms of Service / Telegram).

⁴³⁵ Op. cit. (Community Guidelines / Tiktok).

⁴³⁶ Introduction to the Code / Transparency Centre [Electronic resource]. – URL: <https://disinfocode.eu/introduction-to-the-code/> (accessed: 14.10.2023).

⁴³⁷ 2018 Code of Practice on Disinformation / Shaping Europe's digital future [Electronic resource]. – URL: <https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation> (accessed: 17.01.2023).

specifies cooperation between platforms and supranational institutions with fact-checkers partially included in the process.

However, for all the features of self-regulatory actions by platforms, there is a certain skepticism among observers and researchers about the effectiveness of such initiatives⁴³⁸. However, there is skepticism toward the Code mainly driven by the shortcomings in its text, including the lack of a unified approach in practice, key indicators and methods for determining the achievement of the objective, as well as tools to ensure compliance with obligations⁴³⁹. Despite the increased transparency expressed, in particular, in the reports provided by companies on measures taken, the effectiveness of these measures remains unclear⁴⁴⁰, whereas the commitments remain flexible and vague in favor of corporate interpretation.

The existence of shortcomings was further revealed during the first few years of its implementation in practice⁴⁴¹, particularly in the issues of disproportionality in the implementation of commitments and lack of involvement of other stakeholders. However, even with these shortcomings, the existence of the Code is an indicator of socially responsible dynamics in the fight against disinformation at the platform level, as evidenced by the revision of the Code in 2022⁴⁴², taking into account these shortcomings.

At the same time, the actions of global companies in Europe are currently the first, but not unique, experience of multistakeholder cooperation of digital platforms. For example, in 2021, the Alliance for the Protection of Children in the Digital Environment⁴⁴³ was created in Russia, which took on a narrower range of challenges and a limited audience. However, the focus was nonetheless on malicious content, including fakes and hate speech. Among the practical aspects of the alliance's work, the active

⁴³⁸ Kuczerawy A. Op. cit.; Prokopović A.M. Op. cit.; Katsy D.V. Features of the Implementation of the Common European Code of Practice Regarding Disinformation / D.V. Katsy, L.P. Shmatkova // Actual problems of infotelecommunications in science and education (APINO 2019). – St. Petersburg State University of Telecommunications, 2019. - P. 677-680. (In Russian).

⁴³⁹ The Sounding Board's Unanimous Final Opinion on the So-Called Code of Practice. – Sounding Board of the Multistakeholder Forum on Disinformation Online, 2018.

⁴⁴⁰ Saurwein F. Op. cit. P. 836.

⁴⁴¹ EC Self-assessment Reports / European Commission. – Annual self-assessment reports of signatories to the Code of Practice on Disinformation, 2019; Monti M. The Eu Code of Practice on Disinformation and the Risk of the Privatisation of Censorship / M. Monti // Democracy and Fake News. – Routledge, 2020. – P. 214-225.

⁴⁴² Op. cit. (Introduction to the Code).

⁴⁴³ The largest digital Russian companies join forces to protect children in the digital environment / Rostelecom [Electronic resource]. - URL: <https://www.company.rt.ru/press/news/d460446/> (accessed: 14.10.2023).

development of digital literacy also stands out, pointing to a global trend towards the growing social responsibility of technological companies.

At the same time, despite the practice of inter-actor cooperation, platforms are also turning to other practical tools to combat information influence, which most often means technological instruments to control distributed content. TikTok defines algorithmic moderation and automated flagging of content combined with a fact-checking⁴⁴⁴ process as such means. Other platforms follow similar anti-manipulative information countermeasures, with particular attention to the design and construction of the online platform and the form of the distributed content. Due to the technical details and the connection with the internal architecture of the services, this component of the platform's activities is less public. At the same time, this stage manifests the responsibility of the online platforms for content transfer and display, which might be misinterpreted⁴⁴⁵ due to the inherent polysemantic nature of information. As a result, moderation can raise public concerns about censorship and active interference by platforms in information processes.

For this reason, there is a need for countermeasures associated with greater social and network reflexivity or resilience⁴⁴⁶. Such an understanding of the threat requires a change in hate speech and disinformation positioning as an external phenomenon to the information environment or platform⁴⁴⁷. It means emphasizing the resources for coexistence with information influence through public adaptation. In this context, *the media and civil society organizations* become additional actors in the fight against digital problems. Their response is possible at both global and local levels through fact-checking and digital literacy services⁴⁴⁸, carried out in cooperation with the state and independently.

These actors include, for example, the American educational organization News Literacy Project, which aims to promote practices that develop social and individual awareness and critical thinking⁴⁴⁹. Other organizations offer similar initiatives:

⁴⁴⁴ An update on our work to counter misinformation / Newsroom | TikTok [Electronic resource]. – URL: <https://newsroom.tiktok.com/en-us/an-update-on-our-work-to-counter-misinformation> (accessed: 20.12.2022).

⁴⁴⁵ Paz M.A. Op. cit. P. 6.

⁴⁴⁶ Mikheev E.A. Op. cit.

⁴⁴⁷ Matamoros-Fernández A. Racism, Hate Speech, and Social Media: A Systematic Review and Critique / A. Matamoros-Fernández, J. Farkas // *Television & New Media*. – 2021. – Vol. 22. – № 2. – P. 218.

⁴⁴⁸ Dalkir K. Op. cit.

⁴⁴⁹ About / News Literacy Project [Electronic resource]. – URL: <https://newslit.org/about/> (accessed: 21.12.2022).

Gesellschaft für Medienpädagogik und Kommunikationskultur⁴⁵⁰, Media Education.Fr⁴⁵¹, etc.⁴⁵² Their tasks also include the development of educational standards and teaching materials to promote media literacy. It should be noted that these organizations are predominantly national in scope and do not extend their activities to the territory of other countries, particularly Russia. Their priorities also include the depoliticization of disinformation and hate speech for educational purposes.

In addition to educational initiatives, organizations or media-related fact-checkers play an important role. Such structures also exist at the national level (Faktisk.no⁴⁵³, Pagella Politica⁴⁵⁴, CaptainFact⁴⁵⁵), focusing their activities on verifying data for specific media outlets. Despite the social importance of such organizations, it is worth considering that by shaping the information field regarding individual events and statements or assessing their truthfulness, they participate in the strategic narrative creation, exposing or reinforcing a point of view. This circumstance illustrates certain shortcomings of the approach with stakeholder involvement, whose independence and impartiality must also be verified.

Various research centers that study and systematize existing experiences and promote best management practices through cooperation, seminars, and further research additionally contribute to the fight against information influence. In particular, they include independent organizations (EU DisinfoLab)⁴⁵⁶ and structures established with the support of supranational bodies⁴⁵⁷. Their main task is to develop partnerships between different actors in countering fakes and hatred on the Internet, supporting their actions with scientific expertise. At the same time, they prioritize regional issues aimed at domestic audience.

⁴⁵⁰ GMK / Gesellschaft für Medienpädagogik und Kommunikationskultur [Electronic resource]. – URL: <https://www.gmk-net.de/> (accessed: 21.12.2022).

⁴⁵¹ Accueil / Media Education [Electronic resource]. – URL: <https://mediaeducation.fr/> (accessed: 21.12.2022).

⁴⁵² Media, Information & Digital Literacy Organisations in Europe (old) / EAVI [Electronic resource]. – URL: <https://eavi.eu/media-information-digital-literacy-organisations-europe-old/> (accessed: 21.12.2022).

⁴⁵³ Op. cit. (Faktisk).

⁴⁵⁴ Progetto / Pagella Politica [Electronic resource]. – URL: <https://pagellapolitica.it/progetto> (accessed: 22.12.2022).

⁴⁵⁵ Op. cit. (CaptainFact.io).

⁴⁵⁶ Op. cit. (About us / EU DisinfoLab).

⁴⁵⁷ EDMO at a Glance / EDMO [Electronic resource]. – URL: <https://edmo.eu/edmo-at-a-glance/> (accessed: 22.12.2022).

The aforementioned practices of countering information threats illustrate (through multiple levels of counteraction and a variety of involved stakeholders) the existence of a system of interconnected actors, united by the global nature of the information network. This relationship is universal because, as practice shows, no actor can separately provide necessary actions to reduce information risks for objects of influence. For this reason, recognizing a universal relationship is increasingly becoming part of governance mechanisms that reflect the ideas of a *multistakeholder approach*⁴⁵⁸. This approach emphasizes the interaction of multiple stakeholders and the coordination of efforts, taking into account the actors' interests⁴⁵⁹. At the same time, the term "Multistakeholderism" comes from the sphere of Internet governance⁴⁶⁰, where there is no single dominant regulator, and its place belongs to multiple forums involving government regulators, technical specialists, and other stakeholders.

Based on the diversity of parties involved in combating disinformation and hate speech, the key characteristics of multistakeholderism are the commonality of the problem (meta-problem), collective knowledge, interdependence, and joint governance⁴⁶¹. The conceptual framework outlined makes it possible to identify the main characteristics of the problem under consideration through the experiences and actions of the parties directly or indirectly involved in combating information threats. Such a perspective reveals the distribution of the potential of the network structure and the influence of its parts.

Global or regional institutions more often support multistakeholderism. State actors also take into account its potential, however, the possibility of establishing legally binding practices within their jurisdictions reduces the interest of governments in the delegating or consultative component of multistakeholderism. This is particularly true for countries that seek to fragment the global network. In this context, multistakeholderism can be interpreted as the mere fact that many parties are involved in combating information

⁴⁵⁸ Raymond M. Op. cit.; Antonova S. Op. cit.

⁴⁵⁹ Stadnik I. What Is an International Cybersecurity Regime and How We Can Achieve It? / I. Stadnik // Masaryk University Journal of Law and Technology. – 2017. – Vol. 11. – № 1. – P. 129-154.

⁴⁶⁰ Antonova S. Op. cit.; Epstein D. Multistakeholderism in Praxis: The Case of the Regional and National Internet Governance Forum (IGF) Initiatives / D. Epstein, B.M. Nonnecke // Policy & Internet. – 2016. – Vol. 8. – № 2. – P. 148-173.

⁴⁶¹ Antonova S. Op. cit.

threats (while the hierarchy of involvement may be different) or as a specific governance practice based on the horizontal involvement of various parties (as in the case of the Code of Practice developed by digital platforms).

In the context of governance practices, the ideas of multistakeholderism have a fundamental connection with another governance approach that has become widespread in European and global practice. This approach⁴⁶² relies on the *resilience* concept⁴⁶², which is also associated with the countering practice of systemic threats. In particular, resilience in the context of information threats is actively referred to in the documents of the UN and other international institutions, indicating that resilience, along with multistakeholderism, belongs to the modern governance discourse in the information sphere.

The basis of the resilience concept, having its roots in the scientific understanding of the response of various (biological and social) systems to challenges⁴⁶³, is the relationship between existing *threats* and the *resources* for coping with them. Particular emphasis belongs to the object, sources, and stakeholders⁴⁶⁴, which help to perceive the system's resilience, its resistance to threats, and the assessment of its ability to absorb change. In a more practical context, resilience is the quality of a system, manifesting in the presence or absence of the ability to cope with a challenge⁴⁶⁵. At the same time, resilience itself has characteristics such as *immanence to the system; objectivity, not normativity of the category; focus on system resources, not threats; unified understanding of the system's space without splitting it into external and internal dimensions; multilevel representation of a system consisting of subsystems*⁴⁶⁶. This makes resilience an analytical and governance tool that underpins various threat response models.

Based on the conceptual framework presented, multistakeholderism complements resilience from the standpoint of resources for responding to a threat. Decentralization of

⁴⁶² Gudalov N.N. Semiotics of Resilience in International Relations: The Diversity of Academic and Political Meanings / N.N.Gudalov, D.S.Tulupov // Politeia. – 2018. – № 1 (88). – P. 135-147. (In Russian).

⁴⁶³ Op. cit. (Resilience Concept in the European Union ...).

⁴⁶⁴ Interplays of Sustainability, Resilience, Adaptation and Transformation / J.L. Johnson [et al.] // Handbook of Sustainability and Social Science Research: World Sustainability Series / eds. W. Leal Filho, R.W. Marans, J. Callewaert. – Cham: Springer International Publishing, 2018. – P. 3-25.

⁴⁶⁵ Holling C.S. Resilience and Stability of Ecological Systems / C.S. Holling // Annual Review of Ecology and Systematics. – 1973. – Vol. 4. – P. 1-23.

⁴⁶⁶ Op. cit. (Resilience Concept in the European Union ...). P. 12.

governance contributes to the development of various resilience practices through self-regulation. At the same time, in world practice, resilience (as well as multistakeholderism), contrary to its conceptual foundations, often becomes a normative rhetorical device⁴⁶⁷ that indicates not the essence of ideas and governance methods but a formal, discursive designation or ethical assessment of the actions of the parties. This is especially true of attempts to present a particular practice as the most correct, which exists partly in the practice of Western countries or organizations⁴⁶⁸, where resilience concepts and multistakeholderism appear in the context of liberal democracy.

Regardless of the ethical and normative assessment, resilience provides tools for systems analysis that allow us to differentiate governance practices in terms of their systemic manifestations, meaning that different resilience types should be considered when building models for countering information threats. In particular, resilience researchers D. Chandler, J. Coaffee, and P. Bourbeau have proposed two models based on the relationship between system changeability and incoming challenges⁴⁶⁹. The defining parameter of the system in both classifications is the system's readiness to change to adapt to a new threat. Thus, the model of D. Chandler and J. Coaffee focuses on the resilience levels and the dynamics of transformations, moving from homeostasis to self-regulation and then micro-management. P. Bourbeau selects the degree of distribution of transformation in society as the main variable, focusing on the dynamics from maintaining the status quo to peripheral changes and then to complete renewal. Each model implies a movement from the statics of the system to its transformational dynamics.

This relationship identifies relevant practices for countering disinformation and hate speech through the resource and actor aspects⁴⁷⁰. Based on the willingness to revise

⁴⁶⁷ Kotsur G.V. Normative Dimension in the 'Resilience' concept of the European Union: Internal Contradictions of Neoliberalism / G.V. Kotsur // Bulletin of Perm University. Political Science. – 2019. – Vol. 13. – № 2. – P. 70-77. (In Russian).

⁴⁶⁸ Romanova T.A. Resilience Category in the European Union / T.A. Romanova // Contemporary Europe. – 2017. – № 4 (76). – P. 17-28. (In Russian).

⁴⁶⁹ Introduction: contested paradigms of international resilience // The Routledge Handbook of International Resilience / eds. D. Chandler, J. Coaffee. – Routledge, 2017; Bourbeau P. Resilience, security and world politics / P. Bourbeau // The Routledge Handbook of International Resilience / eds. D. Chandler, J. Coaffee. – Routledge, 2017.

⁴⁷⁰ Romanova T.A. Op. cit. (Disinformation (fake news, propaganda)...); Kolotaev Y.Y. Op. cit. (European Union in the age of post-truth...). P. 73.

the institutional framework and the distribution of power, *paternalistic* and *adaptive* practices are distinguished, examples and illustrations of which correspond to the previously identified experience of different actors.

The paternalistic approach prioritizes methods associated with protective or restrictive measures aimed at all elements of information influence. At the level of the subject, this may mean attempts by states or organizations to prohibit the distribution of harmful content, with subsequent legal liability. This category includes national laws developed in Germany⁴⁷¹, France⁴⁷², and other countries that criminalize misinformation or hate speech online. At the intermediary level, paternalism can come from regulators imposing liability on platforms for dishonest moderation and from the platforms themselves creating filters and algorithms that block certain types of content. In extreme forms, this can take the form of censorship, which is the most characteristic form of paternalism since it is an interference in personal freedoms for security reasons. At the object level, paternalism creates counter-narratives that form an alternative informational environment aimed at reducing the impact of the primary informational threat. From a resilience perspective, paternalism is, thus, correlated with homeostasis or status quo maintenance, i.e., it aims to maintain a static order in the system.

The adaptive model provides alternative response scenarios. In its practical expression, the adaptive approach includes measures aimed at adapting to a threat by qualitatively increasing the system's features, meaning they are durable and flexible. The same logic of action focuses on integrating the threat into the system and finding ways to coexist with it rather than cut it off. Most adaptive measures aim at the potential object of information influence, i.e., at ordinary users or user communities.

The basic means are educational and informational initiatives that change the object from the point of view of the perception of the problem and do not allow it to become a secondary subject. This category of measures includes actions by civil organizations,

⁴⁷¹ Op. cit. (Gesetz zur Verbesserung der Rechtsdurchsetzung...); Startschuss für die Rechtsdurchsetzung in sozialen Netzwerken / Bundesministerium der Justiz und für Verbraucherschutz [Electronic resource]. – URL: https://web.archive.org/web/20171102171222/https://www.bmjv.de/SharedDocs/Artikel/DE/2017/091817_Rechtsdurchsetzung_in_sozialen_Netzwerken.html (accessed: 20.12.2022).

⁴⁷² Op. cit. (LOI n° 2018-1202 du 22 décembre 2018); Op. cit (LOI n° 2020-766 du 24 juin 2020); France's top court rejects core of law targeting online hate speech / Reuters [Electronic resource]. – URL: <https://www.reuters.com/article/us-france-tech-regulation-idUSKBN23P32O> (accessed: 20.12.2022).

media, and platforms to develop critical thinking (i.e., to reduce the impact of cognitive biases), digital literacy, or the ability to respond to online threats. In addition, fact-checking and subsequent labeling of content (creation of disclaimers) become adaptive as long as they are not a form of counter-propaganda. An alternative to regulatory practices for intermediaries and platforms in the form of restrictive laws are ethical codes and best practices, which include optimal scenarios for responding to threats voluntarily accepted by stakeholders. Thus, adaptive measures aim at self-regulation, i.e., they correlate with a peripheral or autopoietic approach and have an active implementation at the level of micro-politics. Due to their long-term nature, they have a high potential to lead to qualitative changes or system upgrades.

From the point of view of the main actors, the two approaches have different degrees of priority. It is important to emphasize that a combination of the two approaches, with one or the other emphasis, can be observed at almost every governance level. At the same time, governmental institutions having power at the national level tend towards a paternalistic response, focusing on criminalizing the actions of the subject of information influence or the inaction of the intermediary. Online platforms, depending on their legal obligations, most actively combine paternalistic and adaptive practices. At the same time, international organizations, fact-checkers, the media, and civil society have the most potential for implementing adaptive measures because they do not have direct power. Therefore, they target the object of influence and help it to adapt to the aggressive information environment. At the same time, most actors are not static in the implementation of their approaches, and, therefore, the combination capacity of paternalism and adaptability is subject to changeable dynamics depending on the delegation and distribution of powers or opportunities.

From the point of view of multistakeholderism, paternalism reflects to a lesser extent the positions of multiple actors. For this reason, the main platforms supporting multistakeholderism also support adaptive measures that involve various stakeholders in governance processes on an equal footing. In practice, this contributes to a deeper understanding of the problem but has a cost in terms of the difficulty of harmonizing multiple positions. For this reason, the most effective form of multistakeholderism is the

parallel involvement of different actors, with each of them having the opportunity to implement their own best practices, which are further taken into account in the development of a collective response mechanism.

However, even in this case, multistakeholderism implies a slow response and can only work if the response to threats is ensured preventively through a combination with paternalistic measures. For this reason, based on the practice of various states, paternalistic measures are most often the primary response to an information threat, whether in the form of regulation, censorship, or counter-narrative. Based on the above practices and models of countering information threats, we can conclude that several variables determine the way the various actors react. The first variable is the choice of the aimed part of the communication environment to which the measures. In this sense, we speak of the subject, the intermediary, or the object of information influence. Depending on the degree of responsibility and mandatory measures, different actors tend to influence only a separate fragment of the information matrix. In particular, media literacy is an object-oriented measure, ethical codes aim at intermediaries, and laws aim at the subject of information influence (or also at intermediaries). At the same time, the pluralism of actors forms a multistakeholder understanding of the processes of dealing with information challenges. It means the need to coordinate the actions of different actors due to the specifics of the digital information space.

In addition, the systemic nature of the measures taken, their focus on transforming or maintaining the status quo, and the emphasis on regulation or grassroots practices act as a second variable. This variable forms adaptive and paternalistic models that characterize the actors' attitudes towards the threat and the choice of response means. The two models illustrate the diversity of practices to counter different types of disinformation and hate speech at various levels of the international system.

The integration of new technical systems, including AI, into the media environment forms a renewed prospective threat to the global community. The initial introduction of

automation and trained algorithms into digital platforms has already transformed the information distribution system and its content. However, the latest forms of AI optimization for the needs of information influence through its malicious use allow the application of highly personalized, enhanced information relevant to specific audiences, becoming in the context of information influence a multiplicative factor.

To date, most AI threats are in the realm of future prediction. However, the rapid development of the technology turns scenarios of potential use of AI for malicious information purposes into a real risk for the global community and national actors. In this regard, it is crucial to consider the already mentioned and to search for new potential threats stemming from AI in the information sphere from both theoretical and practical points of view. The global community is at the starting point of searching for answers to AI challenges, and the initial conceptualization of social and technical vulnerabilities is a priority when trying to form sustainable measures balancing excessive techno-optimism and skepticism.

Concerning already existing best practices in combating information threats, the analysis of international experience illustrated the existence of actions at the global/supranational, national, regional, and local levels. In terms of content, these practices depend on the nature of the actor's perception of the interaction process with the threat. Modern response scenarios have the logic of adapting to systemic effects, using a new challenge as a tool for system evolution. Such a vision relies on the interconnectedness of the system and the threat, thus indicating the absence of the possibility of the system's complete isolation through protective actions. An alternative path is a paternalistic approach that reflects defensive measures aimed at protecting or restoring the existing equilibrium of the system.

Various combinations of these two governance logics indicate a great variability of practices, including regulatory methods, media literacy, content monitoring, moderation, fact-checking, etc. Depending on the actor, certain individual practices or their combination prevail. All of these measures, as well as the cooperation of the actors, constitute a multistakeholder approach to solving problems, expressed in horizontal and vertical cooperation to develop optimal practices, while the resilience concept inspires a

search for more flexible mechanisms to respond to challenges, to coexist with the threat in the context of its permanent transformation.

Conclusion

The results of the study illustrate the actualization of the problem of information influence in the realities of the 21st century. In contrast to previous historical stages, modernity has led to the aggravation of the problem due to qualitative and quantitative changes in the information environment. As a result of the ongoing changes, the digital information space has become overloaded with manipulative and unverified information. The growth of international tensions, bilateral confrontations, and various global crises led to information disorder. In such a situation, socially important information is in a single information flow with disinformation or hate speech. Simultaneously, the choice of preferred information is influenced indirectly by algorithms and policies of large online platforms.

A detailed consideration of theoretical, systemic, and practical aspects led the author of the dissertation to the following conclusions:

First, the modern information environment requires understanding and conceptualization in world politics analysis through a modernized interpretation of the classical philosophical questions of truth in terms of the socially constructed nature of facts. It is essential to understand the formation of the social, power, and normative-cultural consensus in the perception of truth. In this respect, post-truth is a new and promising concept that combines technogenic issues with classical questions about the search for truth within social relations. At the same time, post-truth avoids rationalist reductionism by including emotions and collective sentiments in a single research agenda, suggesting that the emotional response also combines elements of collective consensus and a memorized cognitive pattern potentially subject to manipulation.

Separate issues related to the scientific nature, originality, and reproducibility of the concept constitute, in a sense, a research agenda. Its consistent elaboration demonstrates the concept's viability but does not negate the need for further theoretical conceptualization for its use in the study of international relations. Post-truth, in the interpretation of this work, defines the devaluation of facts and verification as the main principle of the modern public narrative formation. It appears in the space of devaluing

truth, power, and consensus. In such a space, universal ideas become the object of interpretations generated by multiple actors at the individual level through decentralized information dissemination in the 21st century.

Second, a comprehensive study of information influence requires attention to each participant in the information environment. Among them, are the *subject* of influence, who is the message source, the *intermediary* transmitting and interpreting the message, and the *object* of influence, who perceives and interprets the information. Modeling an example of manipulation in the public sphere requires a certain degree of abstraction since there are no mechanisms for taking into account the individual psychological characteristics of each subject or object. In this regard, generalizing approaches are needed, even though they have a certain degree of error, but reflect consensual ideas about collective attitudes, mechanisms, and responses. Among such ideas are the discourses of truth and emotional regimes explaining social responses to informational stimuli at the grassroots and international levels.

At the same time, another substantial aspect is the influence of rational behavioral strategies and their combination with cognitive biases and heuristics at the level of individual decision-making, as well as their impact on the collective response. In addition, the individual level of information influence is of particular importance in modern digital conditions due to the emergence of targeting and personalization technologies based on artificial intelligence. These technologies exploit users' individual cognitive and emotional vulnerabilities, as they might become a target of disinformation or hate speech campaigns.

Third, different types of information threats, including disinformation and hate speech, must be considered in their diversity of form and content. Each of them differs in terms of the subject's intent, scope of objects, rhetorical and visual instruments, patterns of behavior, and the ultimate goal. In particular, one of the most important principles of information threat differentiation is the question of the intent and involvement of the subject of disinformation dissemination. In addition, it is crucial to distinguish between the degree of polarization of political discourse emanating from hate speech and its potential to turn rhetorical techniques into incentives for real action.

Understanding these differences makes it easier to choose a response strategy and helps to avoid politicizing the problem, which blurs basic concepts and makes it impossible to develop best practices. The degree of impact on society and the threat level should determine the appropriate mechanisms, combining long-term and short-term solutions. Misinterpretation of the information threat associated with over- or under-securitization can lead to problems of disproportionate response.

Fourth, the digital environment dynamic forces information threat researchers and practitioners to pay special attention to new technological capabilities development applicable to digital manipulation. Today, one of the central innovations in this area is AI, which forms a whole set of challenges related to internal and external vulnerabilities of information systems and creates actual and potential scenarios for highly optimized information attacks using empathic tools and emotional AI. The latter is especially dangerous in the context of the increasing emotionality of discourse in the era of post-truth politics and the intensification of digital hate speech.

Fifth, global practice shows that, in modern conditions, it is impossible to develop a universal solution to information threats at any single level of the international system. In this regard, most actors appeal to multilateral instruments related to multistakeholder cooperation. A systemic understanding of the threat forces us to look for countermeasures globally and locally. Global solutions are required because of the scale of today's information environment or the role of large media intermediaries in influencing agenda-setting and data flows. The local experience of counteraction plays a critical role because it is impossible to formulate responses to all challenges only through the cooperation of major actors, since the main vulnerability and, at the same time, the target of influence remains collective and individual consciousness, which requires the development of resilience and a culture of coexistence with the threat.

Resources of multi-level and cross-sectoral influence open up broader perspectives for threat management. The distribution of responsibilities, the inclusion of practices of self- and co-regulation, the promotion of civil initiatives, and the cooperation with the scientific community make the emerging system more flexible and open to further adaptation. However, any governance system is highly dependent on the current

ideological and discursive agenda, which determines the extent to which important aspects are taken into account or included in regulation and counteraction, such as non-political disinformation, the formation of personalized information bubbles and their commercial exploitation by platforms, and internal social polarization exacerbated by echo chambers.

Thus, the conclusions show that the hypothesis formulated at the beginning of the work about the renewal of terminology in the research and practical activities to combat information threats is confirmed because the development of the post-truth concept and the spread of the related notions (fake, hate speech, and online disinformation) is the result of an increased interest in the issue of digital manipulation amid international conflicts and crises of the 2010s and early 2020s. It is shown in the growing scholarly interest and the securitization of threats within the framework of various institutions of the world system. Old information threats, modernized in the digital space, require a quick response implemented by global, regional, national, and local actors. Hence, this arouses interest in the problem and creates the need to summarize the theoretical and practical experience, which was accomplished within the framework of this dissertation.

The results presented in the research also allow us to formulate several practical recommendations:

- Effective tackling of disinformation and hate speech at the international and national levels should be based not only and predominantly on exposing the false narrative but on forming a unified critical perception of the information environment, including the deconstruction of narratives through understanding the context and the actors involved. Performing such a task at the individual level has a high cognitive cost. For this reason, building a long-term policy to counter information threats requires significant educational resources to help individuals develop informational competencies through personal and professional improvement. These resources include educational disciplines on critical thinking and media literacy that combine the classical postulates of logic and rhetoric with an understanding of digital infrastructure and advanced trends in this sphere. An alternative or supplement to such disciplines can be the incorporation of digital literacy in the curricula of disciplines in various educational programs.

- Fact-checking resources in the information environment is a necessary component of a unified and integrated information policy. In strategic planning, however, it is important to take into consideration that such resources are not a universal means of combating disinformation. Fact-checking has flaws since an object with certain attitudes (located in an echo chamber) has a low tendency to change its mind and may perceive public verification as an additional argument against the unifying narrative. In addition, fact-checking platforms themselves can become an instrument of strategic communication distribution, which makes them involved in information influence, as seen in the EU (EUvsDisinfo).

- Interaction with platforms in countering information influence should be balanced. The presence of excessive platform competencies carries the risk of algorithmic or structural censorship of online services. The reverse situation of active external regulation can also harm the information infrastructure if redundant platform control limits the adequate performance of the services' activities. For this reason, the optimal scenario for the interaction of society, the state, and platforms draws on a combination of practices of the platforms' voluntary regulation under public control, mediated by the norms and standards of the state or a supranational actor.

- The integrated interaction of various national and regional entities is an up-to-date experience of regional optimization of the global information environment. The cumulative effect of the efforts of a group of national actors through supranational institutions makes it possible to find compromises with transnational information players in a balanced and legitimate way. It leads to the collective protection of public interests and the possibility of developing a single optimal response to the needs and interests of national and local actors. This approach contributes to the consolidation of the information space at the regional level.

The further development of the problem of information influence, disinformation, and various relevant regional and national experiences is of great social importance because today's world faces an exponential growth of communication technologies and volumes of data that can be quantified and instrumentalized in research or information operations. In this regard, there are several important areas for further development of the

problem: 1) the formation of an interdisciplinary analysis of the behavior of information influence subjects in the information environment; 2) the development and analysis of mechanisms for the support of influence objects at the infrastructure level of information platforms; 3) further research on the problem of malicious use of artificial intelligence as a resource for improving information systems for digital manipulation; 4) research on the application of artificial intelligence to detect information manipulation; 5) further analysis of the best practices of the most active actors in the field of combating information threats; 6) formalization of the theoretical basis of the post-truth concept with its subsequent articulation within the framework of various research disciplines.

Information influence as a universal threat to the collective consensus and power legitimation has become an acute, actively articulated international problem in modern society. Post-truth, polarization, and information warfare are central catalysts that have turned the issue of verification and fact-finding into a governance resource of the information society. Unlike other stages of historical development, the realities of the 21st century not only confirm this but also point to the increased precariousness of the entire communication infrastructure. The generally accessible possibility of publicly translating one's position has made public space more open and international relations more inclusive. However, this openness did not only affect the positive dynamics of information development. It made society more susceptible to manipulative influence, making it disproportionately easier to form one's reality in the information flow than to reach a collective truth and consensus.

Sources and Bibliography

Sources

Laws of the Russian Federation:

1. Federal Law of 18.03.2019 No. 31-FZ on Amending Article 15-3 of the Federal Law on Information, Information Technologies and Information Protection [Electronic resource]. - URL: <http://publication.pravo.gov.ru/Document/View/0001201903180031> (accessed: 15.09.2023).
2. Federal Law of 18.03.2019 № 27-FZ on Amending the Code of the Russian Federation on Administrative Offenses. [Electronic resource]. - URL: <http://publication.pravo.gov.ru/Document/View/0001201903180021?index=1> (accessed: 15.09.2023).
3. Federal Law of 01.04.2020 № 100-FZ on Amendments to the Criminal Code of the Russian Federation and Articles 31 and 151 of the Code of Criminal Procedure of the Russian Federation. [Electronic resource]. - URL: <http://publication.pravo.gov.ru/Document/View/0001202004010073> (accessed: 15.09.2023).

Laws of foreign countries:

4. Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz - NetzDG) vom 1. September 2017 / Bundesgesetzblatt, T. I Nr. 61, S. 3352. – 2017.
5. LOI n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information (1) / NOR : MICX1808389L, JORF n°0297 du 23 décembre 2018. – 2018.
6. LOI n° 2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur internet (1) / NOR : JUSX1913052L, JORF n°0156 du 25 juin 2020. – 2020.

Strategies and resolutions of the UN General Assembly:

7. Resolution "Promoting interreligious and intercultural dialogue and tolerance in countering hate speech" A/RES/75/309 / UN General Assembly. – 2021.
8. The UN Strategy and Plan of Action / United Nations [Electronic resource]. – URL: <https://www.un.org/en/hate-speech/un-strategy-and-plan-of-action-on-hate-speech> (accessed: 04.12.2022).
9. United Nations. Resolution adopted by the General Assembly on 24 December 2021 A/RES/76/227 / United Nations – 2022. – P. 1-4.

Materials of the CIS and CSTO:

10. “Deep Fakes”: Concept and Assessment of Potential Impact on National Security. Analytical note. - Moscow: ATC CIS, 2019. - 39 p.
11. Information and Analytical Activity / Anti-Terrorist Center of the CIS Member States [Electronic resource]. - URL: <https://www.cisatc.org/1289/134> (accessed: 09.09.2023).
12. On May 16, a meeting of the leaders of the CSTO member states was held in Moscow, dedicated to the 30th anniversary of the signing of the Collective Security Treaty and the 20th anniversary of the creation of the CSTO. The Statement of the members of the Collective Security Council was adopted (In Russian) [Electronic resource]. – URL: https://odkb-csto.org/news/news_odkb/v-moskve-16-maya-proydet-vstrecha-glav-gosudarstv-chlenov-odkb-posvyashchennaya-30-letiyu-podpisaniya/#loaded (accessed : 05.03.2023).
13. On the signing of the Agreement on Cooperation of the Member States of the Commonwealth of Independent States in the field of information security / Ministry of Digital Development, Communications and Mass Media of the Russian Federation [Electronic resource]. - URL: https://digital.gov.ru/ru/documents/3729/?utm_referrer=https%3a%2f%2fwww.google.com%2f (accessed: 09.09.2023).

Reports and guidelines of international institutions and organizations:

14. EC Self-assessment Reports / European Commission. – Annual self-assessment reports of signatories to the Code of Practice on Disinformation, 2019.
15. Journalism, fake news & disinformation: handbook for journalism education and training (In Russian) / UNESCO Digital Library [Electronic resource]. – URL: <https://unesdoc.unesco.org/ark:/48223/pf0000371785> (accessed: 04.12.2022).
16. Secretary-General UN. Countering disinformation and promotion and protection of human rights and fundamental freedoms / UN. Secretary-General // Publisher: UN. – 2022.
17. Startschuss für die Rechtsdurchsetzung in sozialen Netzwerken / Bundesministerium der Justiz und für Verbraucherschutz [Electronic resource]. – URL: https://web.archive.org/web/20171102171222/https://www.bmjv.de/SharedDocs/Artikel/DE/2017/091817_Rechtsdurchsetzung_in_sozialen_Netzwerken.html.
18. The Sounding Board's Unanimous Final Opinion on the So-Called Code of Practice. – Sounding Board of the Multistakeholder Forum on Disinformation Online, 2018.

Rules and guidelines of digital services and platforms:

19. 2018 Code of Practice on Disinformation / Shaping Europe's digital future [Electronic resource]. – URL: <https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation> (accessed: 17.01.2023).
20. Community Guidelines / Tiktok (in Russian) [Electronic resource]. – URL: <https://www.tiktok.com/community-guidelines> (accessed: 20.12.2022)
21. Introduction to the Code / Transparency Centre [Electronic resource]. – URL: <https://disinfocode.eu/introduction-to-the-code/> (accessed: 14.10.2023).
22. Terms of Service / Telegram [Electronic resource]. – URL: <https://telegram.org/tos/terms-of-service-for-telegram-premium> (accessed: 21.02.2023).

23. YouTube Community Guidelines - How YouTube Works (in Russian) / YouTube [Electronic resource]. – URL: <https://www.youtube.com/howyoutubeworks/policies/community-guidelines/#community-guidelines> (accessed: 21.02.2023).

24. YouTube. Hate speech policy - YouTube Help [Electronic resource]. – URL: <https://support.google.com/youtube/answer/2801939?hl=en> (accessed: 10.11.2021).

Press releases and information from official websites:

25. About / News Literacy Project [Electronic resource]. – URL: <https://newslit.org/about/> (accessed: 21.12.2022).

26. About us / EU DisinfoLab [Electronic resource]. – URL: <https://www.disinfo.eu/about-us/> (accessed: 22.12.2022).

27. Accueil / Media Education [Electronic resource]. – URL: <https://mediaeducation.fr/> (accessed: 21.12.2022).

28. An update on our work to counter misinformation / Newsroom | TikTok [Electronic resource]. – URL: <https://newsroom.tiktok.com/en-us/an-update-on-our-work-to-counter-misinformation> (accessed: 20.12.2022).

29. CaptainFact.io [Electronic resource]. – URL: <https://captainfact.io/> (accessed: 22.12.2022).

30. EDMO at a Glance / EDMO [Electronic resource]. – URL: <https://edmo.eu/edmo-at-a-glance/> (accessed: 22.12.2022).

31. Faktisk. [Electronic resource]. – URL: <https://www.faktisk.no/> (accessed: 22.12.2022).

32. GMK / Gesellschaft für Medienpädagogik und Kommunikationskultur [Electronic resource]. – URL: <https://www.gmk-net.de/> (accessed: 21.12.2022).

33. Hate Speech - Freedom of Expression / Council of Europe [Electronic resource]. – URL: <https://www.coe.int/en/web/freedom-expression/hate-speech> (accessed: 09.12.2022).

34. In “post truth” era, leaders must defend objective, independent media, UN says on Press Freedom Day / UN News [Electronic resource]. – URL: <https://news.un.org/en/story/2017/05/556472-post-truth-era-leaders-must-defend-objective-independent-media-un-says-press> (accessed: 04.12.2022).
35. Inaccurate Information Makes Effective Responses to Complex Health Crises More Difficult, Secretary-General Tells Event on Countering ‘Fake News’ / UN Press [Electronic resource]. – URL: <https://press.un.org/en/2019/sgsm19574.doc.htm> (accessed: 04.12.2022).
36. Infodemic / WHO [Electronic resource]. – URL: <https://www.who.int/health-topics/infodemic> (accessed: 04.12.2022).
37. Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation / WHO [Electronic resource]. – URL: <https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation> (accessed: 04.12.2022).
38. Media, Information & Digital Literacy Organisations in Europe (old) / EAVI [Electronic resource]. – URL: <https://eavi.eu/media-information-digital-literacy-organisations-europe-old/> (accessed: 21.12.2022).
39. Member State Actions / EU DisinfoLab [Electronic resource]. – URL: <https://www.disinfo.eu/resources/member-state-actions/> (accessed: 20.12.2022).
40. Progetto / Pagella Politica [Electronic resource]. – URL: <https://pagellapolitica.it/progetto> (accessed: 22.12.2022).
41. The largest digital Russian companies join forces to protect children in the digital environment / Rostelecom (in Russian) [Electronic resource]. - URL: <https://www.company.rt.ru/press/news/d460446/> (accessed: 14.10.2023).

News articles and reports:

42. British public still believe Vote Leave ‘£350m a week to EU’ myth from Brexit referendum / The Independent [Electronic resource]. – URL: <https://www.independent.co.uk/news/uk/politics/vote-leave-brex-it-lies-eu-pay-money-remain-poll-boris-johnson-a8603646.html> (accessed: 08.02.2023).
43. FAKE ALERT: Old video from Maharashtra viral as ‘massive cyclone’ in Kerala // The Times of India [Electronic resource]. – URL: <https://timesofindia.indiatimes.com/times-fact-check/news/fake-alert-old-video-from-maharashtra-viral-as-massive-cyclone-in-kerala/articleshow/82973398.cms> (accessed: 25.05.2023).
44. Fake videos spread online after Chadian president’s death // The Observers - France 24 [Electronic resource]. – URL: <https://observers.france24.com/en/africa/20210422-fake-videos-spread-online-after-chadian-president-s-death> (accessed: 25.05.2023).
45. France’s top court rejects core of law targeting online hate speech / Reuters [Electronic resource]. – URL: <https://www.reuters.com/article/us-france-tech-regulation-idUSKBN23P32O> (accessed: 20.12.2022).
46. Freedland J. Post-truth politicians such as Donald Trump and Boris Johnson are no joke | Jonathan Freedland [Electronic resource]. – URL: <http://www.theguardian.com/commentisfree/2016/may/13/boris-johnson-donald-trump-post-truth-politician> (accessed: 24.01.2021).
47. Hiley Likely, Retirement Novichok. Experts named the words of the year in Russia (In Russian) [Electronic resource]. – URL: https://www.dp.ru/a/2018/11/23/Hajli_lajkli__pensionnij (accessed: 08.02.2023).
48. Truth or Fake - No, this video doesn’t show thousands of Syrian migrants trying to cross into the EU // France 24 [Electronic resource]. – URL: <https://www.france24.com/en/tv-shows/truth-or-fake/20220923-no-this-video-doesn-t-show-thousands-of-syrian-migrants-trying-to-cross-into-the-eu> (accessed: 25.05.2023).

Bibliography

Monographs:

49. "Post-Truth Politics" and Populism / ed. O.V. Popova. - St. Petersburg: OOO "Skifia-print", 2018. - 216 p. (In Russian).
50. Anderson P. The origins of postmodernity / P. Anderson. – London ; New York: Verso, 1998. – 143 p.
51. Arendt H. Between Past and Future / H. Arendt, J. Kohn. – Annotated edition. – New York: Penguin Classics, 2006. – 320 p.
52. Bakir V. Empathic Media, Emotional AI, and the Optimization of Disinformation / V. Bakir, A. McStay // *Affective Politics of Digital Media*. – Routledge, 2020.
53. Ball J. Post-truth: how bullshit conquered the world. Post-truth / J. Ball. – London: Biteback Publishing, 2017. – 306 p.
54. Baudrillard J. Simulacra and simulation / J. Baudrillard. - Moscow: Postum, 2017. - 240 p. (In Russian).
55. Baudrillard J. The Spirit of Terrorism. There was no Gulf War / J. Baudrillard; transl. A. Kachalov. - Moscow: RIPOL classic, 2016. - 224 p. (In Russian).
56. Baudrillard J. Selected writings / J. Baudrillard, M. Poster. – 2. ed. – Stanford, Calif: Stanford Univ. Press, 2007. – 294 p.
57. Belton I. Cognitive biases and debiasing in intelligence analysis / I. Belton, M. Dhimi // ed. R. Viale. – London: Routledge, 2020.
58. Bleiker R. Postmodernism / R. Bleiker // *Introduction to international relations: Australian perspectives* / ред. R. Devetak, A. Burke, J. George. – Cambridge ; New York: Cambridge University Press, 2007. – P. 86-95.
59. Block D. Post-Truth and Political Discourse / D. Block. – Cham: Springer International Publishing, 2019. – 130 p.

60. Bourbeau P. Resilience, security and world politics / P. Bourbeau // The Routledge Handbook of International Resilience / ред. D. Chandler, J. Coaffee. – Routledge, 2017.
61. Calero Valdez A. Studying Biases in Visualization Research: Framework and Methods / A. Calero Valdez, M. Ziefle, M. Sedlmair // Cognitive Biases in Visualizations / ed. G. Ellis. – Cham: Springer International Publishing, 2018. – P. 13-27.
62. Campbell D. National deconstruction: Violence, identity, and justice in Bosnia / D. Campbell. – Univ of Minnesota Press, 1998. – 328 p.
63. Campbell D. Writing Security: United States Foreign Policy and the Politics of Identity / D. Campbell. – Revised edition. – Minneapolis: University of Minnesota Press, 1998. – 308 p.
64. Cognitive Biases in Visualizations / ed. G. Ellis. – Cham: Springer International Publishing, 2018.
65. Colwell T.M. I.2 Emotives and emotional regimes / T.M. Colwell // Early Modern Emotions. – Routledge, 2016.
66. Cosentino G. Social Media and the Post-Truth World Order: The Global Dynamics of Disinformation. Social Media and the Post-Truth World Order / G. Cosentino. – Cham: Springer International Publishing, 2020. – 147 p.
67. D’Ancona M. Post truth: the new war on truth and how to fight back. Post truth / M. D’Ancona. – London: Ebury Press, 2017. – 167 p.
68. Dalkir K. Navigating Fake News, Alternative Facts, and Misinformation in a Post-Truth World / K. Dalkir; ред. R. Katz. – 1st edition. – Hershey, PA: IGI Global, 2020. – 375 p.
69. Davis E. Post-truth: why we have reached peak bullshit and what we can do about it. Post-truth / E. Davis. – London: Little, Brown, 2017. – 347 p.
70. Durnová A. Understanding Emotions in Post-Factual Politics / A. Durnová. – Cheltenham: Edward Elgar Publishing, 2019. – 192 p.
71. Early Modern Emotions: An Introduction. Early Modern Emotions / ред. S. Broomhall. – London: Routledge, 2016. – 424 p.

72. Ellul J. Propaganda: The Formation of Men's Attitudes. Propaganda / J. Ellul; пер. K. Kellen, J. Lerner. – 1st edition. – New York: Vintage, 1973. – 352 p.
73. Fairclough N. Media Discourse / N. Fairclough. – 1st edition. – London: Bloomsbury Academic, 1995. – 224 p.
74. Farkas J. Post-Truth, Fake News and Democracy: Mapping the Politics of Falsehood. Post-Truth, Fake News and Democracy / J. Farkas, J. Schou. – Routledge, 2019. – 178 p.
75. Fluck M. The concept of truth in international relations theory: critical thought beyond post-positivism. The concept of truth in international relations theory / M. Fluck. – London, United Kingdom: Palgrave Macmillan, 2017. – 247 p.
76. Foucault M. Nietzsche, Genealogy, History / M. Foucault // The Foucault reader / ред. P. Rabinow. – New York: Pantheon Books, 1984. – P. 76-100.
77. Foucault M. Subjectivity and Truth: Lectures at the Collège de France, 1980-1981 : Michel Foucault, Lectures at the Collège de France. Subjectivity and Truth / M. Foucault; eds. F. Ewald, A. Fontana, F. Gros. – London: Palgrave Macmillan UK, 2017. – 331 p.
78. Foucault M. The Government of Self and Others: Lectures at the Collège de France 1982–1983. The Government of Self and Others / M. Foucault; eds. F. Gros, F. Ewald, A. Fontana. – London: Palgrave Macmillan UK, 2010. – 402 p.
79. Frankfurt H.G. On Bullshit / H.G. Frankfurt. – Princeton, NJ: Princeton University Press, 2005. – 81 p.
80. Fuller S. Post-Truth: Knowledge As A Power Game: A Political Theory. Post-truth / S. Fuller; transl. A. Smirnova. – HSE Publishing House, 2021. – 368 p. (In Russian).
81. Fuller S. Post-Truth: Knowledge As A Power Game. Post-Truth / S. Fuller. – New York: Anthem Press, 2018. – 218 p.
82. Giusti S. Democracy and Fake News: Information Manipulation and Post-Truth Politics. Democracy and Fake News / S. Giusti, E. Piras. – Routledge, 2021. – 246 p.

83. Governability and discourse of virtual communities in the conditions of post-truth politics / I.A. Bykov [et al.]; ed. D.S. Martyanov. - St. Petersburg: ElekSys, 2019. (In Russian).
84. Griggs S. Rational Choice in Public Policy: The Theory in Critical Perspective / S. Griggs // Handbook of Public Policy Analysis. – Routledge, 2007.
85. Halpern D. The Psychology of critical thinking / D. Halpern. - St. Petersburg: Peter, 2000. - 512 p. (In Russian).
86. Hate Speech and Polarization in Participatory Society / ред. M. Pérez-Escobar, J.M. Noguera-Vivo Seneviratne K. Myth of 'Free Media' and Fake News in the Post-Truth Era. – Taylor & Francis, 2022.
87. Heuer R.J. Structured Analytic Techniques for Intelligence Analysis / R.J. Heuer, R.H. Pherson. – 1st edition. – Washington, DC: CQ Press, 2010. – 368 p.
88. Holtzhausen D. Strategic Communication: Opportunities and Challenges of the Research Area / D. Holtzhausen, A. Zerfass // The Routledge Handbook of Strategic Communication. – Routledge, 2014.
89. Ilchenko S.N. Fake control, or News that should not be believed: How the media fool us / S.N. Ilchenko. - Phoenix, 2021. - 188 p. (In Russian).
90. Ilyin E.P. Psychology of communication and interpersonal relations / E.P. Ilyin. - St. Petersburg: Peter, 2009. - 576 p. (In Russian).
91. International theory: positivism and beyond / ред. S. Smith, K. Booth, M. Zalewski. – Cambridge ; New York: Cambridge University Press, 1996. – 362 p.
92. Interplays of Sustainability, Resilience, Adaptation and Transformation / J.L. Johnson [et al.] // Handbook of Sustainability and Social Science Research : World Sustainability Series / eds. W. Leal Filho, R.W. Marans, J. Callewaert. – Cham: Springer International Publishing, 2018. – P. 3-25.
93. Introduction: contested paradigms of international resilience // The Routledge Handbook of International Resilience / ред. D. Chandler, J. Coaffee. – Routledge, 2017. – Introduction.

94. Is the EU Disinformation Review Compliant with EU Law? Complaint to the European Ombudsman About the EU Anti-Fake News Initiative / A. Alemanno [et al.]. – Rochester, NY, 2018.
95. Jahankhani H. Cybercrime Profiling and Trend Analysis / H. Jahankhani, A. Al-Nemrat // Intelligence Management: Knowledge Driven Frameworks for Combating Terrorism and Organized Crime / eds. B. Akhgar, S. Yates. – London: Springer, 2011. – P. 181-197.
96. Jamieson K.H. Echo Chamber: Rush Limbaugh and the Conservative Media Establishment. Echo Chamber / K.H. Jamieson, J.N. Cappella. – First Edition. – Oxford ; New York: Oxford University Press, 2008. – 320 p.
97. Kahneman D. Thinking, Fast and Slow / D. Kahneman. – London: Macmillan, 2011. – 512 p.
98. Kalpokas I. A Political Theory of Post-Truth / I. Kalpokas. – Cham: Springer International Publishing, 2019. – 135 p.
99. Karasika V.I., Sternina I.A. Anthology of concepts / Ed. V.I. Karasika, I.A. Sternin. Volume 3. Volgograd: Paradigm, 2006. - 381 p. (In Russian).
100. Keyes R. The post-truth era: dishonesty and deception in contemporary life. The post-truth era / R. Keyes. – 1st ed. – New York: St. Martin's Press, 2004. – 312 p.
101. Klinger U. Network Media Logic: Some Conceptual Considerations / U. Klinger, J. Svensson. – 2015. – Network Media Logic. – P. 23-38.
102. Kuczerawy A. Fighting Online Disinformation: Did the EU Code of Practice Forget about Freedom of Expression? / A. Kuczerawy // Disinformation and Digital Media as a Challenge for Democracy. – Intersentia, 2019. – P. 291-308.
103. Lasswell H.D. The structure and function of communication in society / Lasswell H.D. // The communication of idea / ред. L. Bryson. – New York: Harper and Row, 1948. – P. 37-51.
104. Lasswell H.D. Propaganda technique in the World War / H.D. Lasswell. – New York: P. Smith, 1938.

105. Lim M. Algorithmic Enclaves: Affective Politics and Algorithms in the Neoliberal Social Media Landscape / M. Lim // *Affective Politics of Digital Media*. – Routledge, 2020.
106. Linklater A. The transformation of political community: ethical foundations of the post-Westphalian era. *The transformation of political community* / A. Linklater. – Reprinted. – Cambridge: Polity Press, 2004. – 263 p.
107. Lyotard J.-F. *The Postmodern Condition* / J.-F. Lyotard; transl. N. A. Shmatko. - St. Petersburg: Aleteyya, 1998. - 160 p. (In Russian).
108. *Making sense of international relations theory* / ред. J.A. Sterling-Folker. – Second edition. – Boulder, Colo: Lynne Rienner Publishers, Inc, 2013. – 481 p.
109. McIntyre L. *Post-Truth* / L. McIntyre. – Cambridge, MA: The MIT Press, 2018. – 236 p.
110. McLuhan M. *Understanding Media: The Extensions of Man*. *Understanding Media* / M. McLuhan, L.H. Lapham. – Reprint edition. – Cambridge, Mass: The MIT Press, 1994. – 389 p.
111. McStay A. *Emotional AI: The Rise of Empathic Media*. *Emotional AI* / A. McStay. – First edition. – Los Angeles: SAGE Publications Ltd, 2018. – 248 p.
112. Medeiros J. de. *Conspiracy Theory in Turkey: Politics and Protest in the Age of “Post-Truth”*. *Conspiracy Theory in Turkey* / J. de Medeiros. – Bloomsbury Publishing, 2018. – 222 p.
113. Mengü M.M. *Digital Perception Management* / M.M. Mengü, E. Türk // *Digital Seige*. – Istanbul, Turkey: Istanbul University Press, 2021.
114. *Models and practices of political content management in the online space of modern states in the era of post-truth* / N.A. Ryabchenko [et al.]. - FLINTA, 2021. - 340 p.(In Russian).
115. Monti M. *The Eu Code of Practice on Disinformation and the Risk of the Privatisation of Censorship* / M. Monti // *Democracy and Fake News*. – Routledge, 2020. – P. 214-225.

116. Moore J. *Bush's War For Reelection: Iraq, the White House, and the People. Bush's War For Reelection / J. Moore.* – 1st edition. – Hoboken, N.J: Wiley, 2004. – 400 p.
117. Nielsen R.K. *News media, search engines and social networking SITES as varieties of online gatekeepers / R.K. Nielsen // Rethinking Journalism Again.* – Routledge, 2016.
118. Ocheretyany K.A. *Faket - a unit of digital experience / K.A. Ocheretyany // Philosophical analytics of the digital era.* - St. Petersburg: St. Petersburg State University Publishing House, 2020. - P. 309-324. (In Russian).
119. Pantserev K.A. *The Malicious Use of AI-Based Deepfake Technology as the New Threat to Psychological Security and Political Stability / K.A. Pantserev // Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity : Advanced Sciences and Technologies for Security Applications / eds. H. Jahankhani [et al.].* – Cham: Springer International Publishing, 2020. – P. 37-55.
120. Pariser E. *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think. The Filter Bubble / E. Pariser.* – Penguin, 2011. – 179 p.
121. Pashentsev E. *The malicious use of artificial intelligence against government and political institutions in the psychological area / E. Pashentsev, D. Bazarkina // Regulating Artificial Intelligence in Industry.* – Routledge, 2021.
122. Picard R.W. *Affective Computing / R.W. Picard.* – Cambridge, MA, USA: MIT Press, 1997. – 306 p.
123. Pitruzzella G. *Disinformation and Hate Speech: A European Constitutional Perspective / G. Pitruzzella, O. Pollicino.* – Milano: EGEA Spa - Bocconi University Press, 2020. – 174 p.
124. Plous S. *The Psychology of Judgment and Decision Making / S. Plous .* - Moscow: Filin, 1998. (In Russian).
125. Ponkin I.V. *Non-classical wars / I.V. Ponkin.* – "INFRA-M", 2019. (In Russian).
126. Ponomarev N.F. *Postmodern strategic communications. Post-truth, memes, transmedia / N.F. Ponomarev.* - «KnoRus» Publishing House, 2023. - 174 p. (In Russian).
127. *Post-Truth, Philosophy and Law / eds. A. Condello, T. Andina.* – 1st edition. – Routledge, 2020. – 172 p.

128. Reddy W.M. *The Navigation of Feeling: A Framework for the History of Emotions*. *The Navigation of Feeling* / W.M. Reddy. – Cambridge: Cambridge University Press, 2001.
129. *Research Anthology on Fake News, Political Warfare, and Combatting the Spread of Misinformation* / Information Resources Management Association. – IGI Global, 2020. – 653 p.
130. *Resilience Concept in the European Union: Articulation and its Consequences for Russia* / ed. T.A. Romanova. - St. Petersburg: Publishing House of the SPBU, 2019. (In Russian).
131. Rietdijk N. *Post-Truth, False Balance and Virtuous Gatekeeping* / N. Rietdijk, A. Archer // *Virtues, Democracy, and Online Media: Ethical and Epistemic Issues* / eds. N. Snow, M.S. Vaccarezza. – Routledge, 2021.
132. Rostova N.N. *Philosophy of transmedia: collective monograph* / N.N. Rostov. - Moscow: Prospect, 2022. (In Russian).
133. *Routledge Handbook of Bounded Rationality* / ed. R. Viale. – 1st edition. – London ; New York, NY: Routledge, 2020. – 664 p.
134. Seneviratne K. *Myth of ‘Free Media’ and Fake News in the Post-Truth Era* / K. Seneviratne. – SAGE Publications Pvt Ltd, 2020.
135. Sim S. *Post-Truth, Scepticism & Power* / S. Sim. – Cham: Springer International Publishing, 2019.
136. *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*. / J.A. Tucker [et al.]. – Rochester, NY, 2018.
137. Swartz L. *Disability and Citizenship in the Global South in a Post-truth Era* / L. Swartz // *The Palgrave Handbook of Disability and Citizenship in the Global South* / eds. B. Watermeyer, J. McKenzie, L. Swartz. – Cham: Springer International Publishing, 2019. – P. 57-65.
138. *The Routledge Handbook of Political Epistemology* / eds. M. Hannon, J. de Ridder. – 1st edition. – London ; New York: Routledge, 2021. – 522 p.

139. Tsebelis G. *Nested Games: Rational Choice in Comparative Politics*. *Nested Games / G. Tsebelis*. – University of California Press, 1990. – 300 p.
140. Tsesis A. *Destructive Messages: How Hate Speech Paves the Way for Harmful Social Movements / A. Tsesis*. – New York: New York University Press, 2002.
141. Visvizi A. *Politics and Technology in the Post-Truth Era / A. Visvizi, M.D. Lytras*. – Emerald Group Publishing, 2019. – 320 p.
142. Wahl-Jorgensen K. *Emotions, media and politics : Contemporary political communication / K. Wahl-Jorgensen*. – Cambridge, UK: Polity, 2019.
143. Webster F. *Theories of the Information Society / F. Webster*. – 4th edition. – Abingdon, Oxon: Routledge, 2014. – 404 p.
144. Wight C. *Philosophy of Social Science and International Relations / C. Wight // Handbook of International Relations*. – London: SAGE Publications, 2002. – P. 23-51.
145. Wittgenstein L. *Philosophical investigations / L. Wittgenstein*. – Oxford: Basil Blackwell, 1968. – 250 p.

Research articles and conference proceedings:

146. “Fake News” Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content / M.D. Molina [et al.] // *American Behavioral Scientist*. – 2021. – Vol. 65. – № 2. – P. 180-212.
147. A Survey of Game Theoretic Approaches to Modelling Decision-Making in Information Warfare Scenarios / K. Merrick [et al.] // *Future Internet*. – 2016. – Vol. 8. – № 3.
148. A survey of sentiment analysis in social media / L. Yue [и др.] // *Knowledge and Information Systems*. – 2019. – Vol. 60. – № 2. – P. 617-663.
149. A systematic literature review on disinformation: Toward a unified taxonomical framework / E. Kapantai [et al.] // *New Media & Society*. – 2021. – Vol. 23. – № 5. – P. 1301-1326.

150. Adversarial attacks on a lexical sentiment analysis classifier / G.A. de Oliveira [et al.] // *Computer Communications*. – 2021. – Vol. 174. – P. 154-171.
151. Akimenko G.V. Infomedy: fake news in a pandemic / G.V. Akimenko [et al.] // *Bulletin of social Sciences and Humanities*. – 2021. – Vol. 2. – № 3. – P. 43-52. (In Russian).
152. Alexander J. Disinformation: A Taxonomy / J. Alexander, J. Smith // *IEEE Security & Privacy*. – 2011. – Vol. 9. – № 1. – P. 58-63.
153. Alexeyev A.B. Hate speech and manipulation as peculiarities of the modern political communication / A.B. Alexeyev // *Izvestiya of Saratov University. Sociology. Politology*. – 2021. – Vol. 21. – № 4. – P. 457-463. (In Russian).
154. Antonova S. “Capacity-building” in global Internet governance: The long-term outcomes of “multistakeholderism” / S. Antonova // *Regulation & Governance*. – 2011. – Vol. 5. – № 4. – P. 425-445.
155. Arias-Maldonado M. A Genealogy for Post-Truth Democracies: Philosophy, Affects, Technology / M. Arias-Maldonado // *Communication & Society*. – 2020. – P. 65-78.
156. Bakir V. Fake News and The Economy of Emotions / V. Bakir, A. McStay // *Digital Journalism*. – 2018. – Vol. 6. – № 2. – P. 154-175.
157. Bakir V. Psychological Operations in Digital Political Campaigns: Assessing Cambridge Analytica’s Psychographic Profiling and Targeting / V. Bakir // *Frontiers in Communication*. – 2020. – Vol. 5.
158. Baranov N.A. Soft power in the conditions of post-truth / N.A. Baranov // *Social and Political Research*. – 2018. – № 1. – P. 20-30. (In Russian).
159. Barbu O. Advertising, Microtargeting and Social Media : International Conference on Communication and Education in Knowledge Society / O. Barbu // *Procedia - Social and Behavioral Sciences*. – 2014. – Vol. 163. – P. 44-49.
160. Bazarkina D. Evolution of Approaches to Countering Hybrid Threats in the European Union’s Strategic Planning / D. Bazarkina // *Contemporary Europe*. – 2021. – № 6. – P. 133-143. (In Russian).

161. Bazarkina D. Evolution of the European Union's Approach to Countering Disinformation / D. Bazarkina // *Vlast*. – 2021. – № 6. – P. 130-138. (In Russian).
162. Bazarkina D. Malicious Use of Artificial Intelligence: New Psychological Security Risks in BRICS Countries / D. Bazarkina, E. Pashentsev // *Russia in Global Affairs*. – 2020. – Vol. 18. – № 4. – P. 154-177.
163. Bhargava V.R. Ethics of the Attention Economy: The Problem of Social Media Addiction / V.R. Bhargava, M. Velasquez // *Business Ethics Quarterly*. – 2021. – Vol. 31. – № 3. – P. 321-359.
164. Böcskei B. "Post-truth politics" as the normal state of politics / B. Böcskei // *Tamara Journal of Critical Organisation Inquiry*. – 2017. – Vol. 15. – № 3/4. – P. 257-260.
165. Boler M. The affective politics of the "post-truth" era: Feeling rules and networked subjectivity / M. Boler, E. Davis // *Emotion, Space and Society*. – 2018. – Vol. 27. – P. 75-85.
166. Bontridder N. The role of artificial intelligence in disinformation / N. Bontridder, Y. Poulet // *Data & Policy*. – 2021. – Vol. 3. – P. e32.
167. Braddock R. An Extension of the "Lasswell Formula" / R. Braddock // *Journal of Communication*. – 1958. – Vol. 8. – № 2. – P. 88-93.
168. Brewer P.R. Priming or Framing: Media Influence on Attitudes Toward Foreign Countries / P.R. Brewer, J. Graf, L. Willnat // *Gazette (Leiden, Netherlands)*. – 2003. – Vol. 65. – № 6. – P. 493-508.
169. Brown N.I. Deepfakes and the Weaponization of Disinformation / N.I. Brown // *Virginia Journal of Law & Technology*. – 2020. – Vol. 23. – P. 1.
170. Browne C. Conceptualising the Political Imaginary: An Introduction to the Special Issue / C. Browne, P. Diehl // *Social Epistemology*. – 2019. – Vol. 33. – Conceptualising the Political Imaginary. – P. 1-5.
171. Carrera P. The stratagems of post-truth / Carrera P. // *RLCS. Revista Latina de Comunicación Social*. – 2018. – Vol. 73. – P. 1469-1481.

172. Chakraborty K. A Survey of Sentiment Analysis from Social Media Data / K. Chakraborty, S. Bhattacharyya, R. Bag // *IEEE Transactions on Computational Social Systems*. – 2020. – Vol. 7. – № 2. – P. 450-464.
173. Chernikov M.V. The Truth and the Varsity Concepts in Russian Culture: the Problem of Correlation / M.V. Chernikov // *Polis. Political Studies*. – 1999. № 5. – P. 43-61. (In Russian).
174. Chetty N. Hate speech review in the context of online social networks / N. Chetty, S. Alathur // *Aggression and Violent Behavior*. – 2018. – Vol. 40. – P. 108-118.
175. Chipidza W. The effectiveness of flagging content belonging to prominent individuals: The case of Donald Trump on Twitter / W. Chipidza, J. (Kevin) Yan // *Journal of the Association for Information Science and Technology*. – 2022. – Vol. 73. – № 11. – P. 1641-1658.
176. Chugrov S.V. Post-Truth: Transformation of Political Reality or Self-Destruction of Liberal Democracy? / S.V. Chugrov // *Polis. Political Studies*. – 2017. – № 2. – P. 42-59. (In Russian).
177. Cobbe J. Algorithmic Censorship by Social Platforms: Power and Resistance / J. Cobbe // *Philosophy & Technology*. – 2021. – Vol. 34.
178. Corner J. Fake news, post-truth and media-political change / J. Corner // *Media, Culture & Society*. – 2017. – Vol. 39. – № 7. – P. 1100-1107.
179. Cosentino G. Post-truth politics in the Middle East: the case studies of Syria and Turkey / G. Cosentino, B. Alikasifoglu // *Artnodes*. – 2019. – № 24. – P. 91-100.
180. Crilley R. International relations in the age of 'post-truth' politics / R. Crilley // *International Affairs*. – 2018. – Vol. 94. – № 2. – P. 417-425.
181. Current Response and Management Decisions of the European Union to the COVID-19 Outbreak: A Review / K. Goniewicz [et al.] // *Sustainability*. – 2020. – Vol. 12. – № 9. – P. 3838.
182. Das R. Rethinking Fake News: Disinformation and Ideology during the time of COVID-19 Global Pandemic / R. Das, W. Ahmed // *IIM Kozhikode Society & Management Review*. – 2022. – Vol. 11. – Rethinking Fake News. – № 1. – P. 146-159.

183. De Blasio E. Who Is Responsible for Disinformation? European Approaches to Social Platforms' Accountability in the Post-Truth Era / E. De Blasio, D. Selva // *American Behavioral Scientist*. – 2021. – Vol. 65. – № 6. – P. 825-846.
184. Deception, identity, and security: the game theory of sybil attacks / W. Casey [et al.] // *Communications of the ACM*. – 2018. – Vol. 62. – № 1. – P. 85-93.
185. Deep learning for affective computing: Text-based emotion recognition in decision support / B. Kratzwald [et al.] // *Decision Support Systems*. – 2018. – Vol. 115. – P. 24-35.
186. Defining Strategic Communication / K. Hallahan [и др.] // *International Journal of Strategic Communication*. – 2007. – Vol. 1. – № 1. – P. 3-35.
187. Denilkhanov A.K. Post-truth and the liberal paradigm of postmodernity / A.K. Denilkhanov // *Bulletin of Moscow University. Series 12*. – 2021. – № 6. – P. 34-46. (In Russian).
188. Derman G.S. Perception management in the media / G.S. Derman // *International Journal of Social and Economic Sciences*. – 2021. – Vol. 11. – № 1. – P. 64-78.
189. Dias Oliva T. Content Moderation Technologies: Applying Human Rights Standards to Protect Freedom of Expression / T. Dias Oliva // *Human Rights Law Review*. – 2020. – Vol. 20. – № 4. – P. 607-640.
190. Douai A. Global and Arab Media in the Post-truth Era: Globalization, Authoritarianism and Fake News / A. Douai // *IEMed: Mediterranean Yearbook*. – 2019. – P. 124-132.
191. Durnová A. Unpacking emotional contexts of post-truth. / A. Durnová // *Critical Policy Studies*. – 2019. – Vol. 13. – № 4. – P. 447-450.
192. Effron D.A. The moral psychology of misinformation: Why we excuse dishonesty in a post-truth world / D.A. Effron, B.A. Helgason // *Current Opinion in Psychology*. – 2022. – Vol. 47. – The moral psychology of misinformation. – P. 101375.
193. Eilders C. Media as political actors? Issue focusing and selective emphasis in the German quality press / C. Eilders // *German Politics*. – 2000. – Vol. 9. – Media as political actors? – № 3. – P. 181-206.

194. Emotion AI-Driven Sentiment Analysis: A Survey, Future Research Directions, and Open Issues / P. Chakriswaran [et al.] // *Applied Sciences*. – 2019. – Vol. 9. – № 24. – P. 5462.
195. Entman R.M. Framing Bias: Media in the Distribution of Power / R.M. Entman // *Journal of Communication*. – 2007. – Vol. 57. – № 1. – P. 163-173.
196. Epstein D. Multistakeholderism in Praxis: The Case of the Regional and National Internet Governance Forum (IGF) Initiatives / D. Epstein, B.M. Nonnecke // *Policy & Internet*. – 2016. – Vol. 8. – № 2. – P. 148-173.
197. Eremina N.V. “New Normal”: The End or the Beginning? / N.V. Eremina // *The New Past*. – 2021. – № 3. – P. 198–207. (In Russian).
198. Estellés M. The educational implications of populism, emotions and digital hate speech: A dialogue with scholars from Canada, Chile, Spain, the UK, and the US / M. Estellés, J. Castellví // *Sustainability (Switzerland)*. – 2020. – Vol. 12. – № 15.
199. Evaluating the Effectiveness of Deplatforming as a Moderation Strategy on Twitter / S. Jhaver [и др.] // *Proceedings of the ACM on Human-Computer Interaction*. – 2021. – Vol. 5. – № CSCW2. – P. 381:1-381:30.
200. Experts on the Malicious Use of Artificial Intelligence: Challenges for Political Stability and International Psychological Security / ed. E. Pashentsev. – Moscow: International Center for Social and Political Studies and Consulting, 2020.
201. Fake News Detection Using Sentiment Analysis / B. Bhutani [и др.] // 2019 Twelfth International Conference on Contemporary Computing (IC3). – 2019. – P. 1-5.
202. Falling for fake news: the role of political bias and cognitive ability / E.C. Tandoc [и др.] // *Asian Journal of Communication*. – 2021. – Vol. 31. – Falling for fake news. – № 4. – P. 237-253.
203. Farte G.I. Reactive Public Relations Strategies for Managing Fake News in the Online Environment / G.I. Farte, D.R. Obada // *Postmodern Openings*. – 2018. – Vol. 9. – № 2. – P. 26-44.
204. Fathaigh R.Ó. The perils of legally defining disinformation / R.Ó. Fathaigh, N. Helberger, N. Appelmann // *Internet Policy Review*. – 2021. – Vol. 10. – № 4.

205. Fay B. Theory and metatheory in social science — or, why the philosophy of social science is so hard / B. Fay // *Metaphilosophy*. – 1985. – Vol. 16. – № 2/3. – P. 150-165.
206. Fino A. Defining Hate Speech / A. Fino // *Journal of International Criminal Justice*. – 2020. – Vol. 18. – P. 31-57.
207. Fortuna P. A Survey on Automatic Detection of Hate Speech in Text / P. Fortuna, S. Nunes // *ACM Computing Surveys*. – 2018. – Vol. 51. – № 4. – P. 85:1-85:30.
208. Foster C. Truth as social practice in a digital era: iteration as persuasion / C. Foster // *AI & SOCIETY*. – 2022.
209. Fuller S. The Post-Truth About Philosophy and Rhetoric / S. Fuller // *Philosophy & Rhetoric*. – 2017. – Vol. 50. – № 4. – P. 473-482.
210. Garbuznyak A.Yu. Post Truth Phenomenon: The Devaluation Of Fact In Media Discourse / A. Yu. Garbuznyak // *Knowledge. Understanding. Skill*. – 2019. – № 1. – P. 184-192. (In Russian).
211. Garrido S. Emotional Regimes Reflected in a Popular Ballad: Perspectives on Gender, Love and Protest in ‘Scarborough Fair’ / S. Garrido, J. Davidson // *Musicology Australia*. – 2016. – Vol. 38. – № 1. – P. 65-78.
212. Gaultney I. Political polarization, misinformation, and media literacy / I. Gaultney, T. Sherron, C. Boden // *Journal of Media Literacy Education*. – 2022. – Vol. 14. – № 1. – P. 59-81.
213. Gavra D.P. Category of Strategic Communication: Modern Understanding and Basic Characteristics / D.P. Gavra // *Information Age*. - 2015. - No. 3. - P. 229-233. (In Russian).
214. Gillespie T. Content moderation, AI, and the question of scale / T. Gillespie // *Big Data & Society*. – 2020. – Vol. 7. – № 2.
215. Gorwa R. Algorithmic content moderation: Technical and political challenges in the automation of platform governance / R. Gorwa, R. Binns, C. Katzenbach // *Big Data & Society*. – 2020. – Vol. 7. – № 1.
216. Grech V. Fake news and post-truth pronouncements in general and in early human development / V. Grech // *Early Human Development*. – 2017. – Vol. 115. – P. 118-120.

217. Grishanina T.A. Artificial intelligence in international relations: role and research dimensions / T.A. Grishanina // RSUH/RGGU Bulletin Series "Political Science. History. International Relations". – 2021. – № 4. – P. 10-18. (In Russian).
218. Gromyko A.A. New Reality and European Security / A.A. Gromyko // Contemporary Europe. – 2016. – № 5 (71). – P. 5-10. (In Russian).
219. Grozina N.A. The problem of the source of power and the theme of resistance in Foucault's philosophy / N.A. Grozina // Vestnik NSU. Series: Philosophy. - 2006. - Vol. 4. - No. 2. - P. 153-159. (In Russian).
220. Gudalov N.N. Semiotics of Resilience in International Relations: The Diversity of Academic and Political Meanings / N.N.Gudalov, D.S.Tulupov // Politeia. – 2018. – № 1 (88). – P. 135-147. (In Russian).
221. Gukasov A.V. Systemic and subsystemic levels of foreign policy analysis of modern states / A.V. Gukasov // A new word in science and practice: hypotheses and approbation of research results. - 2014. - No. 12. - P. 83-88. (In Russian).
222. Gunkel D.J. The medium of truth: media studies in the post-truth era / D.J. Gunkel // Review of Communication. – 2019. – Vol. 19. – The medium of truth. – № 4. – P. 309-323.
223. Guo Z. Effect of Disinformation Propagation on Opinion Dynamics: A Game Theoretic Approach / Z. Guo, J. Valinejad, J.-H. Cho // IEEE Transactions on Network Science and Engineering. – 2022. – Vol. 9. – № 5. – P. 3775-3790.
224. Habgood-Coote J. Stop Talking About Fake News! / J. Habgood-Coote // Inquiry: An Interdisciplinary Journal of Philosophy. – 2019. – Vol. 62. – № 9-10. – P. 1033-1065.
225. Hammond-Errey M. Understanding and assessing information influence and foreign interference / M. Hammond-Errey // Journal of Information Warfare. – 2019. – Vol. 18. – №. 1. – P. 1-22.
226. Halliday F. Book review: a “crisis” of international relations? / F. Halliday // International Relations. – 1985. – Vol. 8. – № 4. – P. 407-412.
227. Harsin J. Regimes of Posttruth, Postpolitics, and Attention Economies / J. Harsin // Communication, Culture & Critique. – 2015. – Vol. 8. – № 2. – P. 327-333.

228. Hate speech operationalization: a preliminary examination of hate speech indicators and their structure / J. Papcunová [et al.] // *Complex & Intelligent Systems*. – 2021.
229. Hedling E. Transforming practices of diplomacy: the European External Action Service and digital disinformation / E. Hedling // *International Affairs*. – 2021. – Vol. 97. – № 3. – P. 841-859.
230. Heinrich S. Medical science faces the post-truth era: a plea for the grassroots values of science / S. Heinrich // *Current Opinion in Anaesthesiology*. – 2020. – Vol. 33. – № 2. – P. 198-202.
231. Helberger N. The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power / N. Helberger // *Digital Journalism*. – 2020. – Vol. 8. – № 6. – P. 842-854.
232. Helm R.K. Regulatory Responses to ‘Fake News’ and Freedom of Expression: Normative and Empirical Evaluation / R.K. Helm, H. Nasu // *Human Rights Law Review*. – 2021. – Vol. 21. – № 2. – P. 302-328.
233. Holling C.S. Resilience and Stability of Ecological Systems / C.S. Holling // *Annual Review of Ecology and Systematics*. – 1973. – Vol. 4. – P. 1-23.
234. Huskaj G. State-of-the-art of Scientific Research on Disinformation / G. Huskaj, S. A Axelsson // *European Conference on Cyber Warfare and Security*. – 2023. – Vol. 22. – № 1. – P. 602-609.
235. Hyvönen A.-E. Careless Speech: Conceptualizing Post-Truth Politics¹ / A.-E. Hyvönen // *New Perspectives*. – 2018. – Vol. 26. – Careless Speech. – № 3. – P. 31-55.
236. Ihlebæk K.A. Hostile emotions: An exploratory study of far-right online commenters and their emotional connection to traditional and alternative news media / K.A. Ihlebæk, C.R. Holter // *Journalism*. – 2021. – Vol. 22. – № 5. – P. 1207-1222.
237. Information overload for (bounded) rational agents / E.M. Pothos [et al.] // *Proceedings of the Royal Society B: Biological Sciences*. – 2021. – Vol. 288. – № 1944.
238. Internet, social media and online hate speech. Systematic review / S.A. Castaño-Pulgarín [et al.] // *Aggression and Violent Behavior*. – 2021. – Vol. 58.

239. Ivančik R. Fight Against Disinformation Concerning Coronavirus Pandemas At European Union Level / R. Ivančik // Almanach (Actual Issues in World Economics and Politics). – 2021. – Vol. 16. – № 3. – P. 5-15.
240. Jia F. Misinformation Literature Review: Definitions, Taxonomy, and Models / Jia F. // International Journal of Social Science and Education Research. – 2020. – Vol. 3. – № 12. – P. 85-90.
241. Kahneman D. On the reality of cognitive illusions / D. Kahneman, A. Tversky // Psychological Review. – 1996. – Vol. 103. – P. 582-591.
242. Katsy D.V. Features of the Implementation of the Common European Code of Practice Regarding Disinformation / D.V. Katsy, L.P. Shmatkova // Actual problems of infotelecommunications in science and education (APINO 2019). – St. Petersburg State University of Telecommunications, 2019. - P. 677-680. (In Russian).
243. Khroul V. M. "Post-truth" in a Global Context: Popularity, Heuristic Value and Ontological Essence / V.M. Khroul // Communications. Media. Design. – 2022. – Vol. 7. – № 1. – P. 5-22. (In Russian).
244. Kirchner C. Digital Disinformation: Taxonomy, Impact, Mitigation, and Regulation (Dagstuhl Seminar 21402) / C. Kirchner, F. Roesner // Dagstuhl Reports. – 2022. – Vol. 11. – № 9. – P. 28-44.
245. Kiser E. The Revival of Narrative in Historical Sociology: What Rational Choice Theory can Contribute / E. Kiser // Politics & Society. – 1996. – Vol. 24. – № 3. – P. 249-271.
246. Knoll J. Advertising in social media: a review of empirical evidence / J. Knoll // International Journal of Advertising. – 2016. – Vol. 35. – № 2. – P. 266-300.
247. Kopp C. Classical deception techniques and perception management vs. the four strategies of information warfare / C. Kopp // Conference Proceedings of the 6th Australian Information Warfare and Security Conference. – School of Information Systems, Deakin University, 2005. – P. 81-89.
248. Kopp C. Information-theoretic models of deception: Modelling cooperation and diffusion in populations exposed to “fake news” / C. Kopp, K.B. Korb, B.I. Mills // PLOS ONE. – 2018. – Vol. 13. – № 11.

249. Kopp C. Shannon, Hypergames and Information Warfare / C. Kopp // *Journal of Information Warfare*. – 2003. – Vol. 2. – № 2. – P. 108-118.
250. Korobkov A.D. The Impact of Artificial Intelligence Technologies on International Relations / A.D. Korobkov // *MGIMO Review of International Relations*. – 2021. (In Russian).
251. Koshkarova N.N. On the way to truth leading to lies: the phenomenon of post-truth in modern political communication / N.N. Koshkarova, N.B. Ruzhentseva // *Political Linguistics*. – 2019. – № 1. – P. 50-56. (In Russian).
252. Kostyrev A.G. Post-politics in the Nets of Post-truth / A.G. Kostyrev // *Polis. Political Studies*. – 2021. – P. 64-75. (In Russian).
253. Kotsur G.V. Normative Dimension in the ‘Resilience’ concept of the European Union: Internal Contradictions of Neoliberalism / G.V. Kotsur // *Bulletin of Perm University. Political Science*. – 2019. – Vol. 13. – № 2. – P. 70-77. (In Russian).
254. Koval E.A. Resentment and resentment as reasons for the use of hate rhetoric in media reality / E.A. Koval // *Information Age*. - 2018. - No. 2-1. (In Russian).
255. Kudrna J. The Possibilities of Combating So-Called Disinformation in the Context of the European Union Legal Framework and of Constitutional Guarantees of Freedom of Expression in the European Union Member States / J. Kudrna // *International Comparative Jurisprudence*. – 2022. – Vol. 8. – № 2. – P. 138-151.
256. Lapid Y. The Third Debate: On the Prospects of International Theory in a Post-Positivist Era / Y. Lapid // *International Studies Quarterly*. – 1989. – Vol. 33. – № 3. – P. 235-254.
257. Lasswell H.D. The Theory of Political Propaganda / H.D. Lasswell // *The American Political Science Review*. – 1927. – Vol. 21. – № 3. – P. 627-631.
258. Lemieux V. Leveraging Archival Theory to Develop A Taxonomy of Online Disinformation / V. Lemieux, T.D. Smith // *2018 IEEE International Conference on Big Data (Big Data)*. – 2018. – P. 4420-4426.
259. Lewandowsky S. Beyond Misinformation: Understanding and Coping with the “Post-Truth” Era / S. Lewandowsky, U.K.H. Ecker, J. Cook // *Journal of Applied*

Research in Memory and Cognition. – 2017. – Vol. 6. – Beyond Misinformation. – № 4. – P. 353-369.

260. Lishaev S.A. "Правда" and "истина" (linguistic conceptualization of the world and thematic originality of Russian philosophy) / S.A. Lishaev // Bulletin of the Samara Humanitarian Academy. – 2006. – № 1 (4). – P. 173-209. (In Russian).

261. Loginov N.I. Current Trends in International Research on Cognitive Biases in Decision-Making Processes / N.I. Loginov, A.I. Aleksandrova // Psychology. Journal of the Higher School of Economics. – 2020. – Vol. 17. – № 3. – P. 444-453. (In Russian).

262. Luchinsky Yu.V. "The Government of Lies": The Phenomenon of Steve Tesich's "Post-Truth" / Yu.V. Luchinsky // Information Age. – 2018. – № 2-1. (In Russian).

263. Lyman P. The Domestication of Anger: The Use and Abuse of Anger in Politics / P. Lyman // European Journal of Social Theory. – 2004. – Vol. 7. – № 2. – P. 133-147.

264. Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning. Manipulating Machine Learning / M. Jagielski [и др.] arXiv:1804.00308 [cs]. – arXiv, 2021.

265. Marcinkowski F. Mediatisation of Politics: Reflections on the State of The Concept / F. Marcinkowski // Javnost - The Public. – 2014. – Vol. 21. – № 2. – P. 5-22.

266. Marsden C. Platform values and democratic elections: How can the law regulate digital disinformation? / C. Marsden, T. Meyer, I. Brown // Computer Law & Security Review. – 2020. – Vol. 36. – P. 105373.

267. Martianov V.S. Coronavirus Pandemic and Expert Knowledge Crisis: Reload of Miracle, Mystery and Authority / V.S. Martianov, V.N. Rudenko, L.G. Fishman // Journal of Institutional Studies. – 2022. – Vol. 14. – № 2. – P. 47-58. (In Russian).

268. Martinets Y.A. The Growth of Right-Wing Populism in Modern Europe (According to the Estimates of European Experts) / Y.A. Martinets // RSUH/RGGU Bulletin Series "Political Science. History. International Relations". – 2019. – № 1. – P. 123-138. (In Russian).

269. Martyanov D.S. Emotional public sphere: Polarization of paralinguistic internet discourse / D.S. Martyanov, G.V. Lukyanova // Vestnik Moskovskogo universiteta. Seriya 10. Zhurnalistika. – 2021. – № 2. – P. 25-48. (In Russian).

270. Matamoros-Fernández A. Racism, Hate Speech, and Social Media: A Systematic Review and Critique / A. Matamoros-Fernández, J. Farkas // *Television & New Media*. – 2021. – Vol. 22. – № 2. – P. 205-224.
271. McDermott R. Psychological Underpinnings of Post-Truth in Political Beliefs / R. McDermott // *PS: Political Science & Politics*. – 2019. – Vol. 52. – № 2. – P. 218-222.
272. McGrath L.S. Historiography, affect, and the neurosciences / L.S. McGrath // *History of Psychology*. – 2017. – Vol. 20. – № 2. – P. 129-147.
273. Michelsen N. Post-Truth-Telling in International Relations / N. Michelsen, B. Tallis // *New Perspectives*. – 2018. – Vol. 26. – № 3. – P. 7-15.
274. Mikheev E.A. Disinformation about the new coronavirus infection COVID-19: psychological mechanisms of influence / E.A. Mikheev // *Problems of the National Strategy*. - 2021. - No. 1 (64). (In Russian).
275. Monetizing disinformation in the attention economy: The case of genetically modified organisms (GMOs) / C.D. Ryan [et al.] // *European Management Journal*. – 2020. – Vol. 38. – № 1. – P. 7-18.
276. Morozov A. How “Post-Truth World” Finally Became Hyperstition / A. Morozov // *Philosophy. Journal of the Higher School of Economics*. – 2021. – Vol. 5. – № 3. – P. 287-297. (In Russian)
277. Mustonen-Ollila E. Information Influence in Society’s Information Environment / E. Mustonen-Ollila, M. J. Lehto, J. Heikkonen // *Journal of Information Warfare*. – 2020. – Vol. 19. – № 4. – P. 70-88.
278. Nenadić I. Unpacking the «European approach» to tackling challenges of disinformation and political manipulation / I. Nenadić // *Internet Policy Review*. – 2019. – Vol. 8. – № 4. – P. 1-22.
279. Nerino V. Overcome the fragmentation in online propaganda literature: the role of cultural and cognitive sociology / V. Nerino // *Frontiers of Sociology*. – 2023. – Vol. 8. – P. 1-10.
280. Newman S. Post-truth and the crisis of the political. / S. Newman // *Soft Power*. – 2019. – Vol. 6. – № 2. – P. 90-108.

281. Nosovets S.G. New media: defining the notion / S.G. Nosovets // *Communication Studies*. – 2016. – № 4 (10). – P. 39-47. (In Russian).
282. Noury A. Identity Politics and Populism in Europe / A. Noury, G. Roland // *Annual Review of Political Science*. – 2020. – Vol. 23. – № 1. – P. 421-439.
283. Nye R.A. Review of *The Navigation of Feeling: A Framework for the History of Emotions*, William M. Reddy / R.A. Nye // *The Journal of Modern History*. – 2003. – Vol. 75. – № 4. – P. 920-923.
284. Obadă R. Sharing Fake News about Brands on Social Media: a New Conceptual Model Based on Flow Theory / R. Obadă // *Argumentum. Journal of the Seminar of Discursive Logic, Argumentation Theory and Rhetoric*. – 2019. – Vol. 17. – P. 144-166.
285. Offensive, aggressive, and hate speech analysis: From data-centric to human-centered approach / J. Kocoń [et al.] // *Information Processing & Management*. – 2021. – Vol. 58. – № 5.
286. Orlov M.O. The Diversity of the Digital Environment in a Risk Society/ M.O. Orlov // *Izvestiya of Saratov University. Philosophy. Psychology. Pedagogy*. – 2019. – Vol. 19. – № 2. – P. 155-161. (In Russian).
287. Page B.I. The Mass Media as Political Actors / B.I. Page // *PS: Political Science & Politics*. – 1996. – Vol. 29. – № 1. – P. 20-24.
288. Pan Z. Priming and Media Impact on the Evaluations of the President's Performance / Z. Pan, G.M. Kosicki // *Communication Research*. – 1997. – Vol. 24. – № 1. – P. 3-30.
289. Pashentsev E. Malicious Use of Artificial Intelligence / E. Pashentsev // *European Conference on the Impact of Artificial IntelligenceI and Robotics. ECIAIR 2019*. – UK: Academic Conferences and Publishing International Limited, 2019. – P. 33.
290. Patra R.K. Bibliometric analysis of fake news indexed in Web of Science and Scopus (2001-2020) / R.K. Patra, N. Pandey, D. Sudarsan // *Global Knowledge, Memory and Communication*. – 2022.
291. Pawelec M. Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions / M. Pawelec // *Digital Society*. – 2022. – Vol. 1. – № 2. – P. 19.

292. Paz M.A. Hate Speech: A Systematized Review / M.A. Paz, J. Montero-Díaz, A. Moreno-Delgado // SAGE Open. – 2020. – Vol. 10. – № 4. – P. 2158244020973022.
293. Penzina A.I. Post-truth: Origin and Current Situation / A.I. Penzina // Language and speech on the Internet: Personality, society, communication, culture. - Peoples Friendship University of Russia (RUDN), 2020. – P. 222-228. (In Russian).
294. Petrishev E.V. Destructive informational influence in the context of countermeasures of challenges and threats of national security in modern Russia/ E.V. Petrishev // Central Russian Journal of Social Sciences. – 2017. – Vol. 12. – №. 4. – P. 83-90. (In Russian).
295. Picciotto R. Is evaluation obsolete in a post-truth world? / R. Picciotto // Evaluation and Program Planning. – 2019. – Vol. 73. – P. 88-96.
296. Plantin J.-C. Digital media infrastructures: pipes, platforms, and politics / J.-C. Plantin, A. Punathambekar // Media, Culture & Society. – 2019. – № 2 (41). – P. 163-174.
297. Podrezov M.V. The History of the Study of Political Populism and Its Current Position in Political Science / M.V. Podrezov // Tomsk State University Journal. – 2020. – № 451. – P. 96-101. (In Russian).
298. Ponkin I.V. Artificial Intelligence from the Point of View of Law / I.V. Ponkin, A.I. Redkina // RUDN Journal of Law. – 2018. – Vol. 22. – № 1. – P. 91-109. (In Russian).
299. Pool J. Infodemic, Misinformation and Disinformation in Pandemics: Scientific Landscape and the Road Ahead for Public Health Informatics Research / J. Pool, F. Fatehi, S. Akhlaghpour // Studies in Health Technology and Informatics. – 2021. – Vol. 281. – Infodemic, Misinformation and Disinformation in Pandemics. – P. 764-768.
300. Popescu N. Hybrid tactics: Russia and the West / N. Popescu // European Union Institute for Security Studies, Alert. – 2015. – Vol. 46.
301. Post-truth politics in the modern world / ed. O.V. Popova. – Proceedings of the Russian scientific conference with international participation "Post-truth politics and populism in the modern world" September 22–23, 2017. - St. Petersburg: LLC "Skifia-print", 2017. - 282 p. (In Russian).

302. Prasad A. Anti-science Misinformation and Conspiracies: COVID–19, Post-truth, and Science & Technology Studies (STS) / A. Prasad // *Science, Technology and Society*. – 2022. – Vol. 27. – Anti-science Misinformation and Conspiracies. – № 1. – P. 88-112.
303. Prikhodko M.V. Linguistic model of information war: structural elements and impact levels / M.V. Prikhodko // *Bulletin of the Moscow Region State University. Series: Linguistics*. – 2023. – №. 3. – P. 57-71. (In Russian).
304. Prokopović A.M. The European Approach to Regulating Disinformation / A.M. Prokopović, M. Vujović // *FACTA UNIVERSITATIS - Law and Politics*. – 2020. – Vol. 18. – № 3. – P. 175-183.
305. Prozorov S. Why is there truth? Foucault in the age of post-truth politics / S. Prozorov // *Constellations*. – 2019. – Vol. 26. – № 1. – P. 18-30.
306. Raymond M. Multistakeholderism: anatomy of an inchoate global institution / M. Raymond, L. DeNardis // *International Theory*. – 2015. – Vol. 7. – № 3. – P. 572-616.
307. Robinson M. Cyber warfare: Issues and challenges / M. Robinson, K. Jones, H. Janicke // *Computers & Security*. – 2015. – Vol. 49. – P. 70-94.
308. Rogaleva O.S. New media: evolution of the concept / O.S. Rogaleva, T.V. Shkaiderova // *Herald of Omsk University*. – 2015. – № 1 (75). – P. 222-225. (In Russian).
309. Rogalsky A. I. Expert discussion about the phenomenon of “post-truth”: domestic and foreign discourse / Rogalsky A. I. // *Socio-political processes in a changing world*. - 2020. - No. 20. (In Russian).
310. Rogers R. Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media / R. Rogers // *European Journal of Communication*. – 2020. – Vol. 35. – № 3. – P. 213-229.
311. Romanova E.R. The persuasive power of fake news: the concept of cognitive distortion / E.R. Romanova // *Journal of Dubna State University. Series: «Science of man and society»*. – 2021. – № 3. – P. 110-118. (In Russian).
312. Romanova T.A. Resilience Category in the European Union / T.A. Romanova // *Contemporary Europe*. – 2017. – № 4 (76). – P. 17-28. (In Russian).
313. Rose J. Brexit, Trump, and Post-Truth Politics / J. Rose // *Public Integrity*. – 2017. – Vol. 19. – № 6. – P. 555-558.

314. Rostova N.N. Philosophical Analytics of Idea of the Post-Truth / N.N. Rostova // Christian Reading. – 2018. – № 6. – P. 130-138. (In Russian).
315. Roumate F. Malicious Use of Artificial Intelligence, New Challenges for Diplomacy and International Psychological Security / F. Roumate // Artificial Intelligence and Digital Diplomacy. – 2021. – P. 97-113.
316. Rustamova L.R. Information Influence on the Age of "Post-Truth" and Fake-News / L.R. Rustamova, B.A. Barabash // Political Science Issues. – 2018. – Vol. 8. – № 5. – P. 23-30. (In Russian).
317. Ryabchenko N.A. Political content management: new linguistic units and social practices / N.A. Ryabchenko, V.V. Katermina, O.P. Malysheva // Church, Communication and Culture. – 2019. – Vol. 4. – № 3. – P. 305-322.
318. Ryabchenko N.A. Presidential Campaign in Post-Truth Era: Innovative Digital Technologies of Political Content Management in Social Networks Politics / N.A. Ryabchenko, O.P. Malysheva, A.A. Gnedash // Polis. Political Studies. – 2019. – № 2. – P. 92-106. (In Russian).
319. Salgado S. Online media impact on politics. Views on post-truth politics and postpostmodernism / S. Salgado // International Journal of Media & Cultural Politics. – 2018. – Vol. 14. – № 3. – P. 317-331.
320. Salin A. Post-truth and Its Threats / A. Salin // Russian sociological review. – 2022. – Vol. 21. – № 1. – P. 284-297 (In Russian).
321. Saurwein F. Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe / F. Saurwein, C. Spencer-Smith // Digital Journalism. – 2020. – Vol. 8. – № 6. – P. 820-841.
322. Scheufele D. Framing As a Theory of Media Effects / D. Scheufele // The Journal of Communication. – 1999. – Vol. 49. – P. 103-122.
323. Scheufele D.A. Framing, agenda setting, and priming: The evolution of three media effects models / D.A. Scheufele, D. Tewksbury // Journal of Communication. – 2007. – Vol. 57. – P. 9-20.
324. Sentiment Analysis for Fake News Detection / M.A. Alonso [et al.] // Electronics. – 2021. – Vol. 10. – № 11. – P. 1348.

325. Sepczyńska D. Post-Truth from the Perspective of Hannah Arendt's Political Theory / D. Sepczyńska // *Filozofia*. – 2019. – Vol. 74.
326. Sharikov P. US, Europe and Russia Approaches to the Problem of the Information Policy / P. Sharikov, N. Stepanova // *Contemporary Europe*. – 2019. – № 2 (88). – P. 73-83. (In Russian).
327. Sharoichenko N.D. European Union Policy In The Area Of Countering Soft Power Of Russia In The States Of The Eastern Partnership / N.D. Sharoichenko // *Bulletin of Moscow Region State University*. – 2022. – № 2. (In Russian).
328. Shatin Y.V. Post-Truth as a Rhetorical Phenomenon in Modern Media Space / Y.V. Shatin // *Vestnik NSU. Series: History and Philology*. – 2020. – Vol. 19. – № 6. – P. 250-257.
329. Shchetinina A.V. Hate And Hatred As Forms Of Enmity In Modern Communication: Experience Of Portraiting Lexema And Concept / A.V. Shchetinina, A.S. Semekhina // *Tomsk State Pedagogical University Bulletin*. – 2022. – № 1 (219). – P. 33-42. (In Russian).
330. Shibkova M.O. Populism and Euroscepticism: Correlation of Notions / M.O. Shibkova // *Contemporary Europe*. – 2019. – № 4 (89). – P. 119-129. (In Russian).
331. Shirky C. The Political Power of Social Media: Technology, the Public Sphere, and Political Change / C. Shirky // *Foreign Affairs*. – 2011. – Vol. 90. – The Political Power of Social Media. – № 1. – P. 28-41.
332. Shmatenko L.M. Russian Disinformation Campaigns: What Risk for the European Union and the International Liberal Order? / L.M. Shmatenko // *Academic Workshop The EU and the Crisis of the International Liberal Order: A Systemic Crisis?* – 2019.
333. Siddiquee Md.A. The portrayal of the Rohingya genocide and refugee crisis in the age of post-truth politics / Md.A. Siddiquee // *Asian Journal of Comparative Politics*. – 2020. – Vol. 5. – № 2. – P. 89-103.
334. Simon H.A. Human Nature in Politics: The Dialogue of Psychology with Political Science / H.A. Simon // *The American Political Science Review*. – 1985. – Vol. 79. – № 2. – P. 293-304.

335. Social Bots and the Spread of Disinformation in Social Media: The Challenges of Artificial Intelligence / N. Hajli [et al.] // *British Journal of Management*. – 2022. – Vol. 33. – № 3. – P. 1238-1253.
336. Social Robot Interactions for Social Engineering: Opportunities and Open Issues / A.F. Abate [и др.] // *2020 IEEE Intl Conf (DASC/PiCom/CBDCCom/CyberSciTech)*. – 2020. – P. 539-547.
337. Spiridovich E.A. Problems and the concept of hate speech in the Internet space / E.A. Spiridovich // *Electronic proceedings of young specialist works of the Polotsk State University. Legal Sciences*. - 2021. - Vol. 37. - P. 139-140. (In Russian).
338. Stadnik I. What Is an International Cybersecurity Regime and How We Can Achieve It? / I. Stadnik // *Masaryk University Journal of Law and Technology*. – 2017. – Vol. 11. – № 1. – P. 129-154.
339. Stark L. The Ethics of Emotion in Artificial Intelligence Systems / L. Stark, J. Hoey // *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency : FAccT '21*. – New York, NY, USA: Association for Computing Machinery, 2021. – P. 782-793.
340. Süselbeck J. Sprache und emotionales Gedächtnis. Zur Konstruktion von Gefühlen und Erinnerungen in der Literatur und den Medien / J. Süselbeck // *Emotionen*. – 2019. – P. 282-295.
341. Tandoc E.C. Defining “Fake News”: A typology of scholarly definitions / E.C. Tandoc, Z.W. Lim, R. Ling // *Digital Journalism*. – 2018. – Vol. 6. – № 2. – P. 137-153.
342. Tesich S. A government of lies. / S. Tesich // *Nation*. – 1992. – Vol. 254. – № 1. – P. 12-14.
343. The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review / Y.M. Rocha [et al.] // *Journal of Public Health*. – 2021. – P. 1-10.
344. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. *The Malicious Use of Artificial Intelligence* / M. Brundage [и др.] arXiv: 1802.07228. – 2018.

345. Thirty years of research into hate speech: topics of interest and their evolution / A. Tontodimamma [и др.] // *Scientometrics*. – 2021. – Vol. 126. – № 1. – P. 157-179.
346. Tianyang F. Socio-political turbulence and its communicative consequences / F. Tianyang, N.Yu. Markushina // *Law and Power*. – 2023. – № 2. – P. 4-6. (In Russian).
347. Timmermann W. Counteracting Hate Speech as a Way of Preventing Genocidal Violence / W. Timmermann // *Genocide Studies and Prevention: An International Journal*. – 2008. – Vol. 3. – № 3. – P. 353-374.
348. Tocci N. Resilience and the role of the European Union in the world / N. Tocci // *Contemporary Security Policy*. – 2020. – Vol. 41. – № 2. – P. 176-194.
349. Tomsa D. Public Opinion Polling and Post-truth Politics in Indonesia / D. Tomsa // *Contemporary Southeast Asia*. – 2020. – Vol. 42. – № 1. – P. 1-27.
350. Toward a Comprehensive Model of Fake News: A New Approach to Examine the Creation and Sharing of False Information / A.P. Weiss [et al.] // *Societies*. – 2021. – Vol. 11. – № 3. – P. 82.
351. Tsarik V.S. Countering 'Russian Information Threat' In The European Union Policy After The Ukraine Crisis: Discourse And Institutional Aspects / V.S. Tsarik // *Central Russian Journal of Social Sciences*. – 2020. – Vol. 15. – № 5. (In Russian).
352. Tsarik V.S. Fighting "Russian Disinformation" in the Public Positioning of Western Institutions: An Analysis of the Official Websites of NATO and the European Union / V.S.Tsarik // *Central Russian Journal of Social Sciences*. – 2019. – Vol. 14. – № 6. (In Russian).
353. Tsipursky G. A Psychological Approach to Promoting Truth in Politics: The Pro-Truth Pledge / G. Tsipursky, F. Votta, J.A. Mulick // *Journal of Social and Political Psychology*. – 2018. – Vol. 6. – A Psychological Approach to Promoting Truth in Politics. – № 2. – P. 271-290.
354. Tsvetkova N.A. Digital diplomacy and digital international relations: Challenges and new advantages / N.A. Tsvetkova, A.N. Sytnik, T.A. Grishanina // *Vestnik of Saint Petersburg University. International Relations*. – 2022. – Vol. 15. – № 2. – P. 174–196. (In Russian).

355. Tumanov A.D. The Analysis of Factors of Formation of the Modern System of Political Governance: Global, Regional and Local Levels / A.D. Tumanov // Administrative consulting. – 2021. – Vol. 151. – № 7. – P. 107-115. (In Russian).
356. Tuzovskii I.D. Post-truth as a syndrome of Digital Age: ultimate understanding of the phenomenon and scenarios of the future / I.D. Tuzovskii // Philosophical Thought. – 2020. – № 12. – P. 42 - 60. (In Russian).
357. Tversky A. Judgment under Uncertainty: Heuristics and Biases / A. Tversky, D. Kahneman // Science. – 1974. – Vol. 185. – № 4157. – P. 1124-1131.
358. Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? / P. Barberá [et al.] // Psychological Science. – 2015. – Vol. 26. – № 10. – P. 1531-1542.
359. Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media / B. Zarouali [et al.] // Communication Research. – 2022. – Vol. 49. – № 8. – P. 1066-1091.
360. Vaccari C. Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News / C. Vaccari, A. Chadwick // Social Media + Society. – 2020. – Vol. 6. – № 1.
361. Vasilenko E.N. "Hate Speech": Toward a Definition of a Term / E.N. Vasilenko // Collection of articles of the International Scientific Conference dedicated to the 105th anniversary of Moscow State University named after A.A. Kuleshova. – Mogilev State University named after A.A. Kuleshova, 2019. – P. 126-127. (In Russian).
362. Vérifier S.L. Tackling COVID-19 Disinformation: Internal and External Challenges for the European Union / S.L. Vérifier, C. Bjola, J.A. Koops // The Hague Journal of Diplomacy. – 2020. – Vol. 15. – Tackling COVID-19 Disinformation. – № 4. – P. 569-582.
363. Vese D. Governing Fake News: The Regulation of Social Media and the Right to Freedom of Expression in the Era of Emergency / D. Vese // European Journal of Risk Regulation (EJRR). – 2022. – Vol. 13. – P. 477.

364. Vilson M. The Europeanization of Foreign Policy in the Face of the Russian Disinformation War / M. Vilson // *Sõjateadlane (Estonian Journal of Military Studies)*. – 2016. – № 2. – P. 114-140.
365. Vinogradova T.V. Science and Technology Research (STS) and Post-Truth Politics. (Review) / T.V. Vinogradova // *Social and humanitarian sciences. Domestic and foreign literature. Ser. 8, Science of Science: Abstract Journal*. - 2018. - №. 1. - P. 66-79. (In Russian).
366. Wahl-Jorgensen K. Media coverage of shifting emotional regimes: Donald Trump's angry populism / K. Wahl-Jorgensen // *Media, Culture and Society*. – 2018. – Vol. 40. – № 5. – P. 766-778.
367. Wanasika I. Deception as Strategy: Context and Dynamics / I. Wanasika, T. Adler // *Journal of Managerial Issues*. – 2011. – Vol. 23. – № 3. – P. 364-378.
368. Wang C.-C. Fake News and Related Concepts: Definitions and Recent Research Development / C.-C. Wang // *Contemporary Management Research*. – 2020. – Vol. 16. – Fake News and Related Concepts. – P. 145-174.
369. Wight C. Post-Truth, Postmodernism and Alternative Facts / C. Wight // *New Perspectives*. – 2018. – Vol. 26. – № 3. – P. 17-29.
370. Wijermars M. Sociotechnical imaginaries of algorithmic governance in EU policy on online disinformation and FinTech / M. Wijermars, M. Makhortykh // *New Media & Society*. – 2022. – Vol. 24. – № 4. – P. 942-963.
371. Word-level Textual Adversarial Attacking as Combinatorial Optimization / Y. Zang [и др.] // *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics / arXiv:1910.12196 [cs]*. – 2020. – P. 6066-6080.
372. Yadollahi A. Current State of Text Sentiment Analysis from Opinion to Emotion Mining / A. Yadollahi, A.G. Shahraki, O.R. Zaiane // *ACM Computing Surveys*. – 2017. – Vol. 50. – № 2. – P. 25:1-25:33.
373. Yamey G. Donald Trump: a political determinant of covid-19. / G. Yamey, G. Gonsalves // *BMJ*. – 2020.
374. Yee A. Post-Truth Politics and Fake News in Asia / A. Yee // *Global Asia*. – 2017. – Vol. 12. – P. 66-71.

375. Zarocostas J. How to fight an infodemic / J. Zarocostas // *The Lancet*. – 2020. – Vol. 395. – № 10225. – P. 676.
376. Zemlyanskiy A.V. Propaganda Tasks of the US Media During the Iraq War / A.V. Zemlyanskiy // *Proceedings of Voronezh State University. Series: Philology. Journalism*. – 2022. – № 3. – P. 103-108. (In Russian).
377. Zhuravskaya E. Political Effects of the Internet and Social Media / E. Zhuravskaya, M. Petrova, R. Enikolopov // *Annual Review of Economics*. – 2020. – Vol. 12. – № 1. – P. 415-438.
378. Zolyan S.T. “Fakes” – People Or Texts? Round Table Discussion / S.T. Zolyan // *Terra Linguistica*. – 2021. – Vol. 12. – № 2. – P. 7-32. (In Russian).

Analytical reports and articles:

379. Baines D. Defining misinformation, disinformation and malinformation: An urgent need for clarity during the COVID-19 infodemic : Discussion Papers / D. Baines, R.J.R. Elliott. – Department of Economics, University of Birmingham, 2020.
380. Davies W. The Age of Post-Truth Politics [Electronic resource]. – URL: <https://www.nytimes.com/2016/08/24/opinion/campaign-stops/the-age-of-post-truth-politics.html> (accessed: 24.01.2021).
381. Field of Disinformation, Democratic Processes, and Conflict Prevention: A Scan of the Literature. Field of Disinformation, Democratic Processes, and Conflict Prevention / S. Udupa [et al.] – Social Science Research Council, 2020.
382. Fontevecchia A. The Fallacy Of ‘Deplatforming’ Donald Trump / Forbes [Electronic resource]. – URL: <https://www.forbes.com/sites/afontevecchia/2021/01/20/the-fallacy-of-deplatforming-donald-trump/?sh=b1dcfa411418> (accessed: 10.02.2023).
383. France debates making fake news a crime / Acepresa [Electronic resource]. – URL: <https://www.acepresa.com/english/france-debates-making-fake-news-a-crime/> (accessed: 20.12.2022).

384. Francis Z. Bush, *The Media & Misinformation Surrounding the Iraq War* [Electronic resource]. – URL: <https://www.anewseducation.com/post/iraq-war-misinformation> (accessed: 07.02.2023).
385. Gibbs A. ‘Post-truth’ is the Word of the Year thanks to Brexit, Trump’s election campaigns [Electronic resource]. – URL: <https://www.cnbc.com/2016/11/16/post-truth-is-the-word-of-the-year-thanks-to-brexit-trumps-election-campaigns-says-oxford-dictionaries.html> (accessed: 08.02.2023).
386. Henley J. *Why Vote Leave’s £350m weekly EU cost claim is wrong* : Politics / J. Henley // *The Guardian*. – 2016. – 10 Jun.
387. Lohani A. *Countering Disinformation and Hate Speech Online: Regulation and User Behavioural Change* / ORF [Electronic resource]. – URL: <https://www.orfonline.org/research/countering-disinformation-and-hate-speech-online/> (accessed: 20.12.2022).
388. Scheidt M. *The European Union versus External Disinformation Campaigns in the Midst of Information Warfare: Ready for the Battle?* *College of Europe EU Diplomacy Paper 01/2019* / M. Scheidt. – 2019. – 34 p.
389. Sellars A. *Defining Hate Speech* / A. Sellars. – Berkman Klein Center Research Publication No. 2016-20. – Boston Univ. School of Law, 2016.
390. *Social Media Algorithms: Why You See What You See* / *Georgetown Law Technology Review* [Electronic resource]. – URL: <https://georgetownlawtechreview.org/social-media-algorithms-why-you-see-what-you-see/GLTR-12-2017/> (accessed: 10.02.2023).
391. *The Boris Johnson Brexit Bus Lie of £350m* [Electronic resource]. – URL: <https://www.conversion-uplift.co.uk/brexit/the-boris-johnson-brexit-bus-lie-of-350m/> (accessed: 08.02.2023).
392. Wardle C. *Information disorder: Toward an interdisciplinary framework for research and policymaking* / C. Wardle, H. Derakhshan. – Council of Europe report DGI(2017)09. – 2017.

Dictionaries and encyclopedias:

393. AI and the Future of Disinformation Campaigns / Center for Security and Emerging Technology [Electronic resource]. – URL: <https://cset.georgetown.edu/publication/ai-and-the-future-of-disinformation-campaigns/> (accessed: 12.02.2023).
394. Definition of FACTOID [Electronic resource]. – URL: <https://www.merriam-webster.com/dictionary/factoid> (accessed: 06.02.2023).
395. Djugumanov I. Practical Applications of Advanced Technologies for Enhancing Security and Defense Capabilities: Perspectives and Challenges for the Western Balkans / I. Djugumanov, M. Hadji-Janev. – NATO Science for Peace and Security Series-E: Human and Societal Dynamics. – IOS Press, 2022. – 134 p.
396. Hair J. Religion and Morality (In Russian) / Stanford Philosophical Encyclopedia: Translations of Selected Articles [Electronic resource]. – URL: <https://philosophy.ru/ru/religion-morality/index.php> (date of access: 02/24/2023).
397. Oxford Word of the Year 2016 | Oxford Languages [Electronic resource]. – URL: <https://languages.oup.com/word-of-the-year/2016/> (accessed: 24.01.2021).
398. The role of Social Engineering in evolution of attacks / E. Frumento [и др.]. – DOGANA Consortium, 2016.

Author's papers prepared in the course of writing the dissertation

399. Kolotaev Y. Disinformation in Times of Pandemic: Experience of the EU in Tackling New Digital Threats / Y. Kolotaev // The world after the pandemic: global challenges and prospects for development. Global Development, Iss. 23. IMEMO, 2022 – P. 112-119. (In Russian).
400. Kolotaev Y. Hate Speech in Perception Management Campaigns: New Opportunities of Sentiment Analysis and Affective Computing / Y. Kolotaev // The Palgrave Handbook of Malicious Use of AI and Psychological Security / ed. E. Pashentsev. – Cham: Palgrave Macmillan, 2023. – P. 105–132.
401. Kolotaev Y. Malicious Information Threats of a Post-Covid World in Europe and the European Union // Contemporary Europe. – 2021. – № 7. – P. 143–151. (In Russian).

402. Kolotaev Y. Methods for Countering the Malicious Use of Artificial Intelligence: Northeast Asian Experience and Models / Y. Kolotaev // Information wars. – 2022. – №. 3. – P. 64-68. (In Russian).
403. Kolotaev Y. Political Implications of Hate Speech Digitalization in a Post-Truth Era: Impact on Emotional Regimes in Digital Conflicts / Y. Kolotaev // RUDN Journal of Political Science. - 2022. - Vol. 24. - №3. - P. 517-529. (In Russian).
404. Kolotaev Y. Post-truth through the lens of emotional regimes and emotives // Media in the modern world. 61st Petersburg Readings. St. Petersburg, 2022 - P. 105-106. (In Russian).
405. Kolotaev Y. Prospects and limitations of the scientific use of the post-truth concept / Y. Kolotaev // Future World, Common Efforts, Real Progress. Proceedings of the International scientific-practical conference of young scientists. – Moscow, 2020. – P. 85-89. (In Russian).
406. Kolotaev Y. Sentiment Analysis: Challenges to Psychological Security and Political Stability / Y. Kolotaev// Proceedings of the 3rd European Conference on the Impact of Artificial Intelligence and Robotics ECIAIR 2021, 2021. P. 82-89.
407. Kolotaev Y., Kollnig K. Political influence of online platforms: YouTube's place in European politics / Y. Kolotaev, K. Kollnig // Vestnik of Saint Petersburg University. International Relations. – 2021. – Vol. 14. – №. 2. – P. 225- 240.
408. Kolotaev Y.Y. European Union in the age of post-truth: Developing societal resilience before European Parliament elections 2019 / Y.Y. Kolotaev // Vestnik of Saint Petersburg University. International Relations. – 2020. – Vol. 13, Issue 1. – P. 69–81.
409. Kolotaev Y.Y. Fake news, disinformation and strategic communications as key categories in EU-Russia relations within a post-truth context. External threats and internal challenges / Y. Kolotaev // Russia in the Global World: New Challenges and Opportunities. – 2019. – P. 431-438. (In Russian).
410. Kolotaev Y.Y. Russia's identity formation through the process of its interactions with the European Union within the context of post-truth politics / Y. Kolotaev // Russia in the Global World: New Challenges and Opportunities. – 2018. – P. 101-112. (In Russian).

411. Kolotaev Yu. Yu. Evolution Of The EU's Policy Of Tackling Disinformation. A Multistakeholder Approach / Yu. Yu. Kolotaev // RSUH/RGGU Bulletin Series "Political Science. History. International Relations." – 2023. – № 2. – P. 42-59. (In Russian).
412. Romanova T.A. Disinformation (fake news, propaganda) as a threat to resilience: approaches used in the EU and its member state Lithuania / T. A. Romanova, N. I. Sokolov, Y.Y. Kolotaev // Baltic Region. – 2020. – Vol. 12. № 1. – P. 53-67.

Appendices

Illustration list:

Figure 1 – The relationship of the subject, object, and environment of information influence in post-truth conditions.....	71
Figure 2 – Primary and secondary objects of information influence. Iterative principle.....	72
Figure 3 – The role of the intermediary in the information influence	74
Figure 4 – Forms of disinformation in information influence	83
Figure 5 – “Pyramid of hate”	86
Figure 6 – Subject-object relationships in the spread of hate speech	89
Figure 7 – Forms of information system vulnerability	97

Table list:

Table 1 – Number of publications in Scopus with the keyword “post-truth”.....	42
Table 2 – Number of publications in eLibrary by the keyword “post-truth”	44
Table 3 – Multi-level scenario of psychological attacks using AI.....	99