„dot.pl" - czasopismo Rejestru domeny .pl

Using Critical Realism to analyse Big Data: ontic, epistemic, and ethical assumptions

Małgorzata Stochmal
Uniwersytet Wrocławski, Zakład Socjologii Bezpieczeństwa i Edukacji, Wrocław, Polska
ORCID: https://orcid.org/0000-0002-8385-2844
E-mail: malgorzata.stochmal@uwr.edu.pl

Abstract

Critical realism remains an interesting research program to explore social reality in an in-depth way. A reflection on its creation is intended to alleviate the shortcomings of positivist and post-positivist approaches. The main aim of the article is to present the ontological, epistemological and ethical assumptions of critical realism. These assumptions can be successfully applied when implementing analytical projects related to large data sets. The article is addressed to recipients interested in a critical and realistic vision of the description of reality related to broadly understood big data analytics.

Keywords: Big Data, stratified depth of ontology, the depth of epistemology, Critical realism, transcendent reality

Received: 15.07.2025
Accepted: 07.08.2025
Published: 07.08.2025

Cite this article as:
M. Stochmal, "Using Critical Realism to analyse Big Data: ontic, epistemic, and ethical assumptions"

DOT.PL, no. 1/ 2025, 10.60097/DOTPL/209115

Corresponding author:
Małgorzata Stochmal
Uniwersytet Wrocławski, Zakład Socjologii Bezpieczeństwa i Edukacji, Wrocław, Polska

E-mail:
malgorzata.stochmal@uwr.edu.pl

Introduction

The expansion of human activity in the digital sphere is accelerating, with no signs that its significance will diminish in the foreseeable future. Enthusiasts of Big Data argue that this phenomenon marks a civilizational breakthrough, comparable to the invention of the Internet, the steam engine, or the printing press[1]. We have progressed from storing kilobytes on floppy disks, to megabytes on hard drives, terabytes on disk arrays, and now petabytes in the cloud.[2] This trajectory continues toward even greater volumes of data — zettabytes, yottabytes, and exabytes. Massive data archives are being created to meet the growing demand for digital storage. Many events and interactions that once took place solely in the physical world now occur in digital spaces. As innovative technologies evolve, algorithms — due to the scale of analysis and the complexity of decision-making — not only mediate digital interactions but increasingly shape or even determine the decision-making process itself[3]. Equally striking is the transition from traditional social research to large-scale data digital analytics. The scale and speed of this technological shift are remarkable.

This article aims to provide a metatheoretical reconstruction of the ontological, epistemological, and ethical assumptions of critical realism and to analyze their applicability within the Big Data environment, in response to the question of how these assumptions can be employed to investigate the complexity of digital social reality. The research problem takes the following form: How can the ontological, epistemological, and ethical assumptions of critical realism be applied to the analysis of data in Big Data environments, considering their complex structure and related epistemic and normative challenges?

The techniques currently used in social research are referred to as obsolete, which is associated with some kind of crisis in empirical sociology[4]. Is this really how this

[1] R. Żulicki, *Potencjał Big Data w badaniach społecznych*. „Studia Socjologiczne", 3(266)/2017, p. 175.
[2] Ch. Anderson, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, https://www.wired.com/2008/06/pb-theory/ accesed 31 october 2024.
[3] B. D. Mittelstadt et al., *The ethics of algorithms: Mapping the debate*, „Big Data & Society" 3(2) (2016), p. 2, DOI: 10.1177/2053951716679679.
[4] M. Savage, R. Burrows, *The coming crisis of empirical sociology*, „Sociology" 41(5)/2007, pp. 885–899, DOI: 10.1177/0038038507080443.

situation should be perceived? Definitely not, each type of research creates its own challenges and is conducted in a specific scientific paradigm with the use of adequate analytical techniques. It is true that the analysis of large data sets allows social researchers to delve deeper into the phenomena being explored in a way that is disproportionate to traditional research[5]. The view of reality changes due to the complexity and changeability of its accompanying conditions. We are currently witnessing the fourth paradigm; the individual stages are characterized as follows:

1. Experimental science. Empiricism; describing natural phenomena (pre-Renaissance).
2. Theoretical science. Modelling and generalization (pre-computers).
3. Computational science. Simulation of complex phenomena (pre-Big Data).
4. Exploratory science. Data-intensive; statistical exploration and data mining[6].

Critical realism can also provide the foundation for proper research design, from theoretical assumptions to the construction of the initial model, to the selection of data inputted for analysis, and by how to analyze it and determine the meanings of the results obtained.

The stratified depth of ontology

In critical realism, priority is given to ontology, which ranks above epistemology. Ontology deals with the existence of beings and strong assumptions about their nature. The form of a strongly rooted realistic ontology restores the proper place not only to being, but also to its absence. The fact that we do not perceive a being at some point is not a proof of its non-existence, but it may result from its absence conditioned by the external context. Ontology as a theory of being emphasizes that all beings located in social, cultural, natural, biological or psychological reality exist independently of our any - complete or incomplete, fallible or true — awareness of them. The fact that they exist forces the tacit assumption that they operate in a strictly defined manner (e.g. the law of gravity, gravity, social inequalities, etc.).

[5] K. Krzysztofek, *Big Data Society. Technologie samozapisu i samopokazu*, „Transformacje. Pismo interdyscyplinarne", 1–4(72–55)/2012, pp. 223–257.
[6] R. Kitchin, *Big Data, new epistemologies and paradigm shifts*, „Big Data & Society" 1(1)/2014, p. 3-6, DOI: 10.1177/2053951714528481.

The ontology of being has not been reduced to one dimension, but analyses its coherent structure in depth. Each being should be analyzed in a deathly way, revealing more and more precise characteristics located on its individual layers.

A stratified ontology contains three elements placed in relation to each other: *Domain of Empirical*, d*omain of Actual* and d*omain of Real* [7]. A visualization of this concept is presented in Diagram 1.

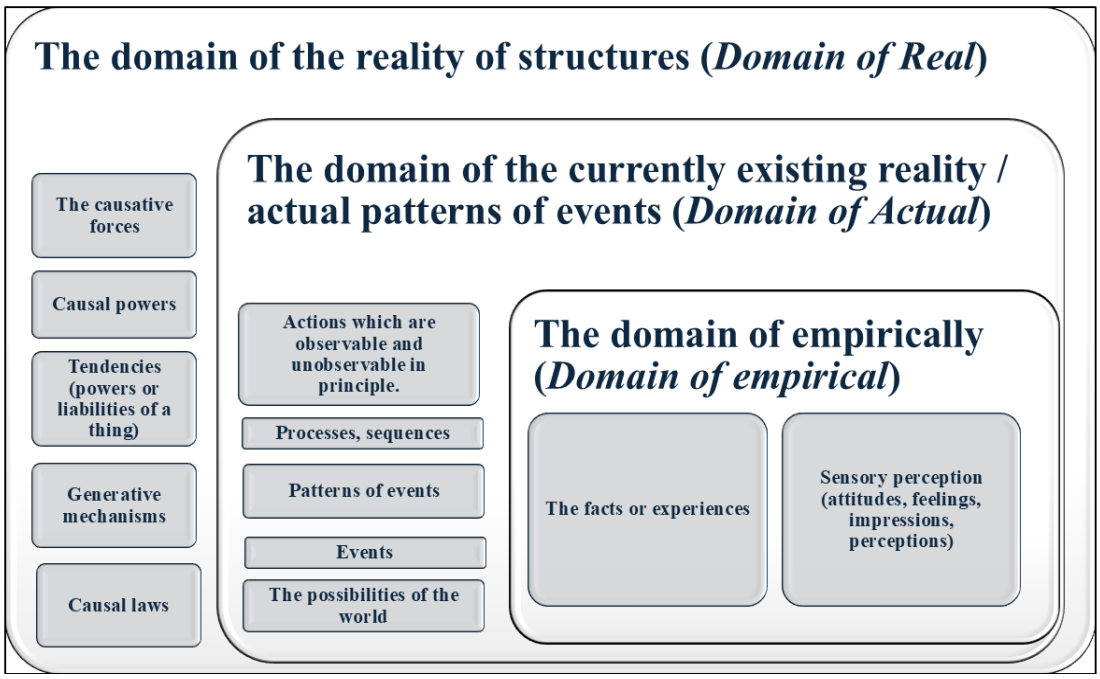

Diagram 1: *Ontological in-depth structures of reality in critical realism*
Source: Translated from anonymous

Roy Bhaskar proposed a stratified and differentiated model of reality that reveals the structured complexity of the ontological realm. According to his theory, reality consists of three analytically distinct but interrelated domains: the empirical, the actual, and the real. These domains are emergent, meaning that each arises from the conditions of the one beneath it while retaining a degree of relative autonomy. They are also temporally displaced, which is particularly significant in the context of scientific inquiry, as research is rarely conducted in real time but often retrospectively, based on empirical traces of past events.

[7]R. Bhaskar, *Enlightened Common Sense: The Philosophy of Critical Realism*, In Introduction Mervyn Hartwig, Pub. Routledge, Abingdon 2016, s. 6-7.

Bhaskar highlighted what he termed the "emergent powers of emergence"—the idea that new causal powers arise at higher levels of ontological complexity. The stratified ontology he proposed does not merely reflect the natural order in a linear fashion; rather, it captures the deep relationality and necessary co-existence of entities and mechanisms across levels of reality. This ontological depth aligns with the complex, layered nature of social and natural phenomena, making it a powerful framework for scientific explanation.

An important consideration, often overlooked in data-driven environments, is the potential mismatch between the purposes for which users generate or provide data and the purposes for which such data are subsequently collected, categorized, and analysed. From a critical realist perspective, this discrepancy underscores the ontological stratification of social reality: while users act within the empirical domain, expressing intentions or needs, the systems that process their input operate on assumptions situated at the actual or even the real domain. Recognizing this misalignment is essential for uncovering the generative mechanisms that underlie the data structures and for avoiding epistemic fallacies rooted in surface-level correlations.

Within this framework, the empirical domain constitutes the most immediately accessible level of reality—what is experienced through observation and sense perception. This includes *facts*, *events*, *sensations*, and *practices* as they appear to human consciousness. However, Bhaskar warns against empirical reductionism: the idea that what is experienced exhausts what is real. Experiences should not be taken as the sole basis for scientific generalization, but rather as the surface expressions of deeper mechanisms.

It is crucial to distinguish between what is *experienced* and what actually *occurs*. The world exists independently of our perception of it, and many events or mechanisms may remain unexperienced yet real. For instance, rainfall is a real event governed by causal mechanisms. However, due to space-time conditions, some people will experience it directly (by getting wet), while others will not—despite the event occurring in the same ontological reality. This example illustrates the transfactual nature of reality in critical realism: mechanisms can operate even when their effects are not empirically manifest.

The domain of the actual refers to the level of reality in which events and processes occur, regardless of whether they are experienced or observed. It occupies the intermediate stratum within Bhaskar's stratified ontology, situated between the domain of the real (generative mechanisms) and the domain of the empirical (experiences). The term "actual" is significant because it denotes the ontological status of events that have taken place or are taking place, independent of their empirical registration. Bhaskar himself states that "(t)he intelligibility of scientific change (and criticism) and scientific education thus presupposes the ontological independence of the objects of experience from the objects of which they are the experiences"[8]. This underscores the necessity of distinguishing between what exists and occurs and what is perceived or known.

Events in the actual domain are caused by real mechanisms, but they may or may not be empirically apprehended. This domain thus includes patterns of events, practices, and interactions that are shaped by underlying structures and mechanisms but are not reducible to immediate experience. It is within this domain that the conditions are set for empirical experience to be possible — including the availability of resources, conceptual frameworks, social practices, and relational contexts. While analytically distinct, the actual and empirical domains are causally linked: empirical observations are always contingent expressions of what has actually occurred, and both are grounded in the generative mechanisms located in the real domain. These events are situated at the middle level of the stratified ontology of reality. They comprise the conditions and configurations—including resources, concepts, practices, and relationships—that make experience in the empirical domain possible.

We thus arrive at the final layer - the domain of real structures, which constitutes the level at which generative mechanisms and causal laws emerge. These mechanisms produce phenomena that manifest as events in the actual domain and may be experienced empirically. Bhaskar defines the function of generative mechanisms as follows:

[8]R. Bhaskar, *A Realist Theory of Science*, with an Introduction by Mervyn Hartwig, Pub. Routledge, London and New York 2008, p. 21.

> “The real basis of causal laws are provided by the generative mechanisms of nature. Such generative mechanisms are, it is argued, nothing other than the ways of acting of things. And causal laws must be analysed as their tendencies. Tendencies may be regarded as powers or liabilities of a thing which may be exercised without being manifest in any particular outcome” (Bhaskar 2008:3)[9].

In the domain of real structures, causal forces are disclosed that activate the powers of generative mechanisms, which, in response to specific conditions, lead subjects to act in particular ways within the empirical domain. Generative mechanisms thus have the potential to produce events, which subsequently become manifest through perceived experience. These mechanisms generate event sequences that support the formulation of causal laws, expressing regularities observable under specific conditions.

While maintaining the ontological rigor of identifying entities embedded in social reality, this bottom-up and iterative process of determining their nature must be conducted carefully. However, the outcomes of such inquiry should not be distorted by the subjectivism of the researcher. The ontological depth structures of reality, as conceived in critical realism, may thus be described as a distinct form of ontological morphology.

Epistemic meanders

The domain in which scientific laws emerge is epistemology, understood as the sphere concerned with the generation and production of scientific knowledge. A central assumption of critical realism is epistemic relativism, which affirms the inherent fallibility of human knowledge and the historically and socially conditioned nature of the truths that individuals formulate.

“Epistemic realism means that all our claims are socially and historically conditioned. Our judgments are determined by circumstances, by what we know at the time and by binding criteria of judgment. For this reason, among other things, our judgments are always error prone. Epistemic relativism then means that each of us is in a situation from

[9]Ibidem, p. 3.

which we see the world in a slightly different way. Our experiences of the world are different"[10].

This fallibility arises from the spatio-temporal context in which knowledge is produced—it is generated at a specific historical moment and, over time, may be revised, distorted, or falsified. The social production of knowledge is burdened with factors that may unintentionally obscure or misrepresent the truth. Bhaskar characterizes science as "a process-in-motion", asserting that "knowledge must be viewed as the produced means of production, and science as a constant social activity in a continuous process of transformation." Furthermore, knowledge is always mediated by concepts, language, history, and social constructs that emerge within particular contexts. The production of knowledge takes place within open social systems, which are inherently dynamic and structurally contingent.

Alongside epistemic relativism and the recognition of ontological depth, attention must also be given to the rationality of judgments. This rationality pertains to the capacity to evaluate competing theoretical frameworks and to adopt the one that most adequately explains the phenomenon in question. Scientific knowledge advances through rational discourse, wherein scholars present evidence and engage in reasoned argumentation. The explanation that most convincingly accounts for the object of inquiry prevails. It must also be acknowledged that all knowledge production is, to some extent, mediated by the researcher's perspective, their interpretation of reality, and their conceptual resources.

Bhaskar distinguishes between two fundamental dimensions of knowledge: the transitive and the intransitive. Transitive objects of knowledge serve as the initial means of production through which intransitive knowledge is formulated. The transitive dimension encompasses the raw materials of science—that is, conceptual and methodological constructs shaped by the prevailing scientific paradigms of a given historical moment. These are inherently changeable, reflecting the critical character of science as a self-transformative enterprise that occasionally yields breakthrough insights into the world. As Bhaskar explains: "Scientists try to discover the reasons for things and events, patterns and processes, sequences and structures. To understand how they do

[10]M. S. Archer, A. Collier, D. V. Porpora, *Transcendencja. Realizm krytyczny i Bóg*, trans. and Introduction Artur Wysocki, epilog Krzysztof Wielecki, Pub. UKSW, Warszawa 2021, p.55.

so one needs both a concept of the transitive process of knowledge-production and a concept of the intransitive objects of the knowledge they produce: the real mechanisms that generate the actual phenomena of the world, including as a special case our perceptions of them"[11].

Objects of transitory knowledge connect the adopted "facts and theories, paradigms and models, methods and techniques of inquiry available to a particular scientific school or worker[12]. This is the natural sequence of things, before any content is included in the field of science, it reveals itself in social spaces and constitutes the basis for formulated generalizations. "In this way social products, antecedently established knowledges capable of functioning as the transitive objects of new knowledges, are used to explore the unknown (but knowable) intransitive structure of the world"[13]. Knowledge in the transitive dimension differs from knowledge embedded in the intransitive dimension.

Defining intransitive objects of knowledge, Bhaskar states that "the intransitive objects of knowledge are in general invariant to our knowledge of them: they are the real things and structures, mechanisms and processes, events and possibilities of the world; and for the most part they are quite independent of us. They are not unknowable, because as a matter of fact quite a bit is known about them[14]. This citation is crucial in the epistemic realm. First, the language used to describe intransitive objects of knowledge is relatively stable; these are truths ultimately expressed within the framework of a given paradigm (as in the case of quantum physics and discoveries nominated for the Nobel Prize — knowledge that extends and deepens classical physics). Second, these truths are formulated in terms of causal laws. The intransitive dimension of science enables a coherent understanding of reality. The law of gravity, for example, belongs to the domain of intransitive knowledge — and knowing it, along with the observable fact that falling objects accelerate, no reasonable person would jump from the tenth floor.

---

[11]R. Bhaskar, *A Realist Theory of Science*, op. Cit., p. 52.
[12]R. Bhaskar, *Realistyczna teoria nauki*, trans. Katarzyna Zahorodna, [in:] M. Sikora (red.), *Realizm wobec wyzwań antyrealizmu. Multidyscyplinarny przegląd stanowisk*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2011, p. 237.
[13]Ibidem, p. 239.
[14]R. Bhaskar, *A Realist Theory of Science*, op. Cit., p. 12.

Formulated in terms of true knowledge or scientific laws, the findings pertain to the operation of generative mechanisms. As Bhaskar states: "The goal of science, however, is to generate knowledge of the mechanisms of creating phenomena in nature that come together to generate a real and continuous variation of world phenomena"[15]. These mechanisms, disclosed through in-depth reflection, function independently of human activity. The assumption of the existence of beings independent of human perception constitutes a transcendental condition necessary for the advancement of scientific inquiry. In the context of critical realism, causal laws remain operative even in the absence of observable events, as they possess an intransitive character. "By saying that the objects of discovery and scientific research are 'intransitive' I mean the indication that they exist independently of any human activity; and by saying that they are 'structured' I mean that they are separate from the patterns of events that occur"[16].

The goal of epistemology is to understand how the world works—specifically, to identify the underlying forces, processes, or mechanisms that generate observable effects or events within the empirical domain. Within its foundational assumptions, critical realism draws a clear distinction between ontology and epistemology, yet acknowledges their analytical interdependence and mutual co-constitution[17]. Epistemic complexity allows for the formulation of knowledge that is not necessarily tied to direct ontological experience. Both ontological and epistemological assumptions are essential for defining social reality and making scientific knowledge possible.

The emergence of generative mechanisms and their causal powers occurs through the application of retroductive reasoning combined with abductive inference. Bhaskar defines these processes as follows:

“Abduction involves redescription or recontextualization, most usually (in CR research) in terms of a characteristic causal mechanism or process which serves to explain it. Retroduction involves imagining a model of a mechanism, which, if it were real,

[15]R. Bhaskar, *Realistyczna teoria nauki*, op. Cit., p. 231.
[16]R. Bhaskar, *A Realist Theory of Science*, op. Cit., 249.
[17] A. K. Albert, J. S. Brundage, P. Sweet i F. Vandenberghe, *Towards a critical realist epistemology?*, „Journal for the Theory of Social Behaviour” 50(3)/2020, p. 358, DOI: 10.1111/jtsb.12248.

would account for the phenomenon in question. (These two can often shade into each other: there is only a relative difference between them)”[18].

Abductive inference involves a reconfiguration of premises that leads to plausible conclusions, based on existing knowledge about the phenomenon under investigation. It offers an alternative to the limitations of purely inductive or deductive reasoning by proposing a logically grounded approach to problem-solving.

In contrast, the retroductive logic of model discovery entails imagining the necessary conditions for the existence of particular elements or phenomena. As Bhaskar describes it: "A thought operation involving a reconstruction of the basic conditions for anything to be what it is, or, to put it differently, it is by reasoning we can obtain knowledge of what properties are required for a phenomenon to exist." The transfactual or transcendental argument represents a form of retroduction, through which one seeks these essential properties beyond what is immediately given”[19].

Ethical assumption

The metatheory of critical realism provides a framework for addressing ethical questions from the standpoint of moral realism. Bhaskar articulates this position in the phase of Dialectical Critical Realism, where the pursuit of alethic truth, freedom, and justice is treated as intrinsically valuable. „Bhaskar’s argument for the universality of morality is a component of his dialectical critical realist ethics; this is a moral realist and ethical naturalist position that seeks to ground moral theory in an understanding of reality”[20]. This ontologically transcendental ethical dimension persists regardless of whether individuals explicitly recognize moral values in their everyday agency. The existence of social reality-itself ontologically real and constitutive of human subjectivity-renders moral agency immanently embedded in human action. Morality thus permeates real structures and activates moral powers capable of transforming the social world. It

[18]R. Bhaskar, *Foreword*, [In:] P. K. Edwards, J. O’Mahoney i S. Vincent (red.), *Studying Organizations Using Critical Realism. A Practical Guide*, Oxford University Press, Oxford 2014, p. VII.
[19]D. Berth, M. Ekstrom, L. Jakobsen, J. Ch. Karlsson, *Explaining Society: An Introduction to Critical Realism in the Social Sciences*, Pub. Routledge, London and New York 2001, p. 206.
[20]S. Ash, *Explaining Morality: Critical Realism and Moral Questions*, Pub. Routledge, London and New York 2022, p. 30.

constitutes the basis of the onto-axiological agency of every subject[21]. Morality constitutes the foundation of the onto-axiological agency of every subject.

Morality is embedded in the intransitive dimension of knowledge. As Steve Ash observes: "This can be understood as stating that moralities are transitive, but they have an intransitive object – intrinsic value[22]. In this view, the dogma of Weberian axiological neutrality is challenged. Values may resonate within subjective human agency, functioning not merely as normative declarations of what ought to be done, but as a transformative force shaping human action.

Critical and realistic data analysis in the area of Big Data

The Big Data environment is a collective term referring to large volumes of data, along with the capabilities for their storage, processing, visualization, and the development of conclusions based on them. There exists a strong relational interdependence among the components of the Big Data environment. In the relevant literature, it is often described as a socio-technological phenomenon, encompassing both social contexts-such as communities, collectives, and groups-and technological infrastructures[23]. Technologies associated with Big Data support a data-driven approach, which involves continuous responsiveness to real-time data outputs and the adaptation of practical actions accordingly[24]. The resulting high-velocity data streams, commonly referred to as Big Data, provide valuable material for research, whether for scientific, business, or other purposes.

When conducting research in the Big Data environment, it is important to recognize the numerous advantages offered by a critical realist approach. At the ontological level, data is embedded within a relatively broad social context and exhibits a high degree of complexity. Before contextualizing such data, it is essential to properly identify its ontological status. Understanding data begins with disclosing its sources of

[21]J. Mariański (red.), *Leksykon socjologii moralności. Podstawy – teorie – badania – perspektywy*, Zakład Wydawniczy NOMOS, Kraków 2015, p. 975.
[22]S. Ash, *Explaining Morality: Critical, op. Cit.* p. 42.
[23]D. Boyd, K. Crawford, *Critical questions for big data*, „Information, Communication & Society" 15(5)/2012, p. 663, DOI: 10.1080/1369118X.2012.678878.
[24]D. Stephenson, *Big data, nauka o danych i AI bez tajemnic. Podejmuj lepsze decyzje i rozwijaj swój biznes!*, transl. Wojciech Bombik, Pub. Helion, Gliwice 2020, p. 72.

origin and the properties ascribed to it. The emergent entities are initially expressed in natural language[25]. This requires a deeper reflection on their identity across the three layers of reality. One should proceed from the level of data (Domain of the Empirical), through the Domain of the Actual, and culminate in identifying the causal mechanisms within the Domain of the Real that generate observable regularities. This process involves the emergence of ontological entities and their intelligibility as epistemologically accessible phenomena[26].

The epistemic dimension, subordinate only to the ontic dimension, is essential in light of the scientific progress already achieved within the relevant field. When engaging in the pursuit of new knowledge, it is necessary to critically assess existing contributions. A thorough analysis of prior scientific achievements helps delineate the intransitive dimension of knowledge, which serves as the foundation for producing, modifying, or refining further knowledge. The researcher must possess a comprehensive understanding of the subject matter within their domain of inquiry. Attention should also be paid to international scholarship. In the era of globalization, access to knowledge across disciplines is merely a few clicks away.

An illustrative example is the issue of fear, which has become a frequent subject of inquiry among Big Data researchers, particularly through techniques such as sentiment analysis. From an ontological perspective, any form of existential uncertainty contributes to the dynamics of social change. Such change may occur in positive, negative or neutral forms.

Elemer Hankiss argues that existential security, which includes the fear of threats among other factors, has been a major driving force behind civilizational transformations. On the one hand, fear has influenced the structural dimension of society by becoming embedded in institutions responsible for ensuring broadly understood security. On the other hand, it has shaped the symbolic and protective sphere, including myths and religions, systems of values and beliefs, ideas and scientific theories, moral and practical

[25]V. Lytvyn, V. Vysotska, O. Veres, *Ontology of Big Data Analytics*, „MEST Journal" 6(1)/2018, p. 54, DOI: 10.12709/mest.06.06.01.06.
[26]Anonymous PWN, Warszawa 2021, p. 100 and next.

norms of behaviour, as well as a wide array of everyday rituals and trivialities[27]. Fear is undoubtedly an emotion that resonates with human action, either motivating individuals to act or causing them to withdraw. These responses can be broadly categorized into three general lines of action, although in practice more variations may exist. The first includes positive actions aimed at overcoming fear. The second involves inaction or the delegation of responsibility for creating safe living conditions to designated institutions. The third encompasses neutral responses, reflecting a reluctance or refusal to engage in any form of action.

An interesting feature of Big Data analysis is the involvement of experts from various scientific disciplines and professional backgrounds in the execution of such projects. Members of these interdisciplinary teams must develop a shared vocabulary that ensures mutual understanding. Another important aspect is that data on the observed phenomenon is collected in its entirety, which makes it possible to systematically monitor the phenomenon and identify recurring patterns.

„The data are not subject to every ontological framing possible, or every form of data-mining technique in the hope that they reveal some hidden truth. Rather, theoretically informed decisions are made as to how best to tackle a data set such that it will reveal information which will be of potential interest and is worthy of further research"[28].

When searching for generative mechanisms, it is necessary to define their components and the strength of the interactions between them. This allows for the identification of patterns or trends emerging from the data. It is important to take into account multiple competing causal forces and powers, and to select the most plausible ones through the application of retrodictive reasoning.

Critical realism established ontic and epistemic assumptions, which were later supplemented by axiomatic assumptions and those pertaining to the transcendental and even theological dimensions of reality. One of the main challenges in conducting research within the digital sphere is the maintenance of ethical standards, particularly those

[27]E. Hankiss, *Fears and Symbols: An Introduction to the Study of Western Civilization*, Pub. CEU Press, Budapest 2001, p. 1-2.
[28] R. Kitchin, *Big Data, new epistemologies*, op. Cit., p. 6.

already established in studies of the experienced, material world. Among the ethical concerns are issues related to data anonymization, methods of data collection, and the tracking of individuals without their informed consent. Determining the potential and actual impact of an algorithm in ethical terms is complex for several reasons. Assessing the influence of human subjectivity in the design and configuration of algorithms often requires the analysis of long-term, multi-user development processes[29].

Final conclusions

Researchers engaged in large-scale data analytics can successfully embed their analytical projects within the framework of critical realism. This perspective addresses the limitations of positivist and post-positivist paradigms, offering a more robust ontological and epistemological foundation for inquiry. Critical realism is increasingly recognized as a valuable approach to understanding social reality, including its digital dimensions[30]. Critical realists focus on mapping the ontological character of social reality, incorporating not only its structural and relational aspects but also its epistemological and axiological dimensions, particularly as they manifest in virtual contexts. The core assumptions of critical realism can thus be effectively applied to research on virtual reality, regardless of the specific subject of investigation. As a research programme, critical realism continues to provide a powerful framework for in-depth exploration of social reality. Reflecting on its theoretical foundations helps to overcome the epistemological and methodological limitations of earlier scientific approaches. It is also worth mentioning the interesting position of Krzysztof Wielecki:

"Virtual space is becoming a new type of social practice in which social, economic, political, cultural, and personal relations are produced, transformed, reproduced, or mediated, along with people's mentalities, values, imaginaries, and belief systems. In this way, life unfolds within a certain split between two dramatically emerging and competing realities at the turn of historical epochs: the virtual reality-largely dominated by mass culture, which appears more alluring, attractive, significant, interesting, and seemingly more true and real-and the other reality, which, by contrast,

[29]Mittelstadt B. D. et all., *The ethics of algorithms, op cit*., p. 2.
[30]K. Wielecki, *Prawda socjologiczna i realizm krytyczny*, „Rocznik Filozoficzny Ignatianum", XXIV(1)/2018, p. 67.

lacks these aforementioned qualities, but is also free from the shortcomings of the former”[31].

Virtual space has become a dominant arena of activity for many individuals. When studying virtual reality, it is essential to consider the specific conditions under which users engage with it. Researchers working within the framework of critical realism possess the conceptual tools necessary to analyse virtual phenomena, including those related to Big Data. Such sources can disclose new dimensions of knowledge that are mediated through transhumanist processes. Therefore, virtual spaces and the reality of the social world should be examined through a hybrid analytical lens, one that integrates both digital and social ontologies.

In conclusion, several key concepts developed throughout the article are crucial for understanding how critical realism can be applied to Big Data research. First, the stratified ontology (empirical–actual–real) provides a robust framework for distinguishing between observable data, the events they represent, and the underlying generative mechanisms. This stratification helps prevent empirical reductionism and supports deeper explanatory models.

Second, the distinction between transitive and intransitive dimensions of knowledge clarifies how scientific understanding is both socially produced and aimed at uncovering reality that exists independently of our cognition. In Big Data environments, this duality invites reflection on how algorithmic knowledge is constructed and how it can obscure or reveal deeper causal patterns.

Third, moral realism and ethical agency remind us that data analytics is not value-neutral. Ethical considerations-such as the anonymization of data, informed consent, and algorithmic accountability-are grounded in ontologically real moral structures that shape, and are shaped by, human action.

---

[31]K. Wielecki, *Kultura versus ~~Kultura~~ masowa. Podmiotowość i quasi-kultura w nibyspołeczeństwie*, Pub. Narodowe Centrum Kultury, Warszawa 2024, p. 595. Original text: Wirtualna przestrzeń staje się nowym rodzajem praktyki, w której wytwarzane, transformowane, reprodukowane lub zapośredniczane są stosunki społeczne, gospodarcze, polityczne, kulturowe, osobiste wraz z mentalnością, wartościami, wyobrażeniami, wierzeniami ludzi. W ten sposób życie przebiega w jakimś rozdwojeniu pomiędzy dwiema dramatycznie wyłaniającymi się i konkurującymi na przełomie epok rzeczywistościami: wirtualną – w ogromnej mierze zdominowaną przez kulturę masową, która wydaje się bardziej powabna, atrakcyjna, ważniejsza, ciekawsza, prawdziwsza i realna, i tą drugą, która – przeciwnie – nie posiada tych wcześniej wymienionych zalet, ale też brak jej wad tamtej.

Finally, the logic of discovery in Big Data should rely not only on induction or correlation but also on abduction and retroduction. These inferential modes allow researchers to formulate hypotheses about hidden mechanisms and necessary conditions for observed data patterns, moving beyond surface-level trends toward deeper causal explanations.

Together, these concepts form a critical-realist foundation for navigating the epistemic and ethical challenges of data-intensive research in the digital age.

References

Albert A. K., Brundage J. S., Sweet P., Vandenberghe F., *Towards a critical realist epistemology?*, „Journal for the Theory of Social Behaviour" 50(3)/2020, p. 358, DOI: 10.1111/jtsb.12248.

Archer M. S., Collier A. Porpora D.V., *Transcendencja. Realizm krytyczny i Bóg*, transl. and Introduction: Artur Wysocki, epilog Krzysztof Wielecki, Pub. UKSW, Warszawa 2021.

Ash S., *Explaining Morality: Critical Realism and Moral Questions*, Pub. Routledge, London and New York 2022.

Berth D., Ekstrom M., Jakobsen L., Karlsson J. Ch., *Explaining Society: An Introduction to Critical Realism in the Social Sciences*, Pub. Routledge, London and New York 2001.

Bhaskar R., *A Realist Theory of Science*, with an Introduction by Mervyn Hartwig, Pub. Routledge, London and New York 2008.

BhaskarR., *Realistyczna teoria nauki*, trans. Katarzyna Zahorodna, [in:] M. Sikora (red.), *Realizm wobec wyzwań antyrealizmu. Multidyscyplinarny przegląd stanowisk*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2011.

BhaskarR. *Foreword*, [in:] P. K. Edwards, J. O'Mahoney i S. Vincent (red.), *Studying Organizations Using Critical Realism. A Practical Guide*, Oxford University Press, Oxford 2014, p. VII.

Bhaskar R., *Enlightened Common Sense: The Philosophy of Critical Realism*, In Introduction Mervyn Hartwig, Pub. Routledge, Abingdon 2016.

Biłobran Cz., *Społeczne konstruowanie operatorów dominacji nad zagrożeniami występującymi lokalnie na przykładzie powiatu nyskiego* https://repozytorium.uni.wroc.pl/dlibra/publication/144569/edition/133980/spoleczne-konstruowanie-operatorow-dominacji-nad-zagrozeniami-wystepujacymi-lokalnie-na-przykladzie-powiatu-nyskiego-bilobran-czeslaw

Boyd, Crawford K., *Critical questions for big data*, „Information, Communication & Society" 15(5)/2012, p. 663, DOI: 10.1080/1369118X.2012.678878.

Hankiss E., *Fears and Symbols: An Introduction to the Study of Western Civilization*, Pub. CEU Press, Budapest 2001.

Kitchin R., *Big Data, new epistemologies and paradigm shifts*, „Big Data & Society" 1(1)/2014, p. 3-6, DOI: 10.1177/2053951714528481.

Krzysztofek K., *Big Data Society. Technologie samozapisu i samopokazu*, „Transformacje. Pismo interdyscyplinarne", 1–4(72–55)/2012, pp. 223–257.

Anderson Ch., *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, https://www.wired.com/2008/06/pb-theory/ accesed 31 October 2024.

Lytvyn V., Vysotska V., Veres O., *Ontology of Big Data Analytics*, „MEST Journal" 6(1)/2018, p. 54, DOI: 10.12709/mest.06.06.01.06.

Mariański J. (red.), *Leksykon socjologii moralności. Podstawy – teorie – badania – perspektywy*, Zakład Wydawniczy NOMOS, Kraków 2015, p. 975.

Mingers J., Mutch A., Willcocks L., *Critical Realism in Information Systems Research*, „IS Quarterly" 37(3)/2013, p. 797, http://irep.ntu.ac.uk/id/eprint/20232/1/216160_300.pdf (accesed: 2 X 2024).

Mittelstadt B. D., Allo P. , Taddeo M., Wachter S. , Floridi L., *The ethics of algorithms: Mapping the debate*, „Big Data & Society" 3(2) (2016), p. 2, DOI: 10.1177/2053951716679679.

Porpora D.V, *A reflection on critical realism and ethics*, „Journal of Critical Realism" 18(3)/2019, p. 274, DOI: 10.1080/14767430.2019.1618064.

Savage M., Burrows R., *The coming crisis of empirical sociology*, „Sociology" 41(5)/2007, pp. 885–899, DOI: 10.1177/0038038507080443.

Stephenson D., *Big data, nauka o danych i AI bez tajemnic. Podejmuj lepsze decyzje i rozwijaj swój biznes!*, transl. Wojciech Bombik, Pub. Helion, Gliwice 2020.

Anonymous Pub. PWN, Warszawa 2021.

Anonymous „Uniwersyteckie Czasopismo Socjologiczne" DOI: 10.21697/ucs.2023.31.1.02.

Wielecki K., *Prawda socjologiczna i realizm krytyczny*, „Rocznik Filozoficzny Ignatianum", XXIV(1)/2018, pp. 45–70.

Wielecki K., *Kultura versus Kultura masowa. Podmiotowość i quasi-kultura w nibyspołeczeństwie*, Pub. Narodowe Centrum Kultury, Warszawa 2024.

Żulicki R, *Potencjał Big Data w badaniach społecznych*, „Studia Socjologiczne" 3(266)/2017, pp. 175–207.