

FILOZOFIA I NAUKA

Studia filozoficzne i interdyscyplinarne

Tom 13, część 2

2025

Instytut Filozofii i Socjologii
Polskiej Akademii Nauk

ISSN 2300-4711
ISSN 2545-1936 (ONLINE)

RADA REDAKCYJNA

Oryśya Bila – Uniwersytet Katolicki, Lwów, Ukraina

Peter Bołtuć – University of Illinois, Springfield, USA

Francesco Coniglione – Uniwersytet w Katanii, Włochy

Adrián Figueroa – Universidad Autónoma de San Luis Potosí, Meksyk

Marek Łagosz — Instytut Filozofii, Uniwersytet Wrocławski

Sofia Miguens — Uniwersytet w Porto, Portugalia

Zbysław Muszyński — Instytut Filozofii, Uniwersytet Marii Curie-Skłodowskiej

Zdzisława Piątek — Instytut Filozofii, Uniwersytet Jagielloński

Aldona Pobjewska — Instytut Filozofii, Uniwersytet Łódzki

María José Frápolli Sanz — Uniwersytet w Grenadzie, Hiszpania

Rui Sampaio da Silva — Uniwersytet Azorów, Portugalia

Pavlo Sodomora – Narodowy Uniwersytet Medyczny im. Danylo Halytsky'ego, Lwów;
Lwowski Uniwersytet Narodowy im. Ivana Franka, Lwów, Ukraina

Peter Sykora — Centrum Bioetyki, UCM University, Trnava, Słowacja

Carlos José B. Tiago de Oliveira — Centro de Filosofia das Ciências da Universidade de Lisboa, Lisbona, Portugalia

Andrew Targowski — Western Michigan University, USA

Barbara Tuchańska — Instytut Filozofii, Uniwersytet Łódzki

Paweł Zeidler — Instytut Filozofii, Uniwersytet im. Adama Mickiewicza, Poznań

FILOZOFIA I NAUKA

Studia filozoficzne i interdyscyplinarne

Tom 13, część 2, 2025



Instytut Filozofii i Socjologii Polskiej Akademii Nauk

ZESPÓŁ REDAKCYJNY

Piotr (Peter) Bołtuć (redaktor naczelny)

Stanisław Czerniak (socjologia wiedzy, filozofia społeczna);
Alicja Kubica (filozofia przyrody);
Włodzimierz Ługowski (filozofia przyrody);
Mariusz Mazurek (filozofia nauki, obsługa strony internetowej);
Miloš Taliga (epistemologia, metodologia, filozofia nauki);
Jordan Zlatev (semiotyka kognitywna)

Skład komputerowy: Mariusz Mazurek

Kontakty

Adres redakcji: Instytut Filozofii i Socjologii Polskiej Akademii Nauk,
pokój 104, ul. Nowy Świat 72, 00-330 Warszawa
Numer telefonu: 603 160 505
adres elektroniczny: filozofiainauka@ifispan.edu.pl
Strona internetowa: <https://filozofiainauka.studiafilozoficzne.edu.pl>

Dostęp

Egzemplarze wersji papierowej — zamówienie z adresem należy składać na elektroniczny adres redakcji.

Czasopismo w wersji elektronicznej jest wydawane w trybie open access: streszczenia wszystkich opublikowanych w czasopiśmie tekstów są dostępne na stronie internetowej czasopisma. Pełne teksty są udostępniane 1) na stronie internetowej czasopisma równocześnie z wersją papierową, 2) na platformie EBSCO, 3) na platformie cyfrowej PAN – Czytelnia Czasopism, a abstrakty – są publikowane w CEJSH (The Central European Journal of Social Sciences and Humanities).

SPIS TREŚCI

Od Redakcji – <i>Editorial</i>	5
Howard Schneider, Piotr (Peter) Boltuć – <i>The Social Halting Problem and the Need for Mitigation of Deceptive Alignment</i>	7
Andrew Targowski – <i>The Quantum Information Theory: Towards the Paradigm of Cognition, the Theory of Everything and Philosophy</i>	19
Stanisław Buda – <i>Wszechświat jako system izolowany</i>	59
Aneta Rumak – <i>Koncepcja stylu myślowego Ludwika Flecka a rozumienie zjawiska wielokulturowości</i>	79
Jacek Gurczyński – <i>Manipulacja emocjami i dane afektywne w epoce cyfrowej...</i>	99
Sławomir Czetwertyński, Jakub Marcinkowski – <i>O heurystykach badawczych w nauce instytucjonalnej w perspektywie nauk ekonomicznych</i>	113

Od Redakcji

Rozpoczynamy nowy rozdział w piśmie „*Filozofia i Nauka. Studia filozoficzne i interdyscyplinarne*”. Pani Profesor Małgorzata Czarnocka, założycielka pisma i jego długoletnia redaktor naczelna, zrezygnowała z pełnionej funkcji, a ja miałem zaszczyt wygrać konkurs na tę funkcję.

Obecny numer rozpoczynają dwa artykuły w języku angielskim. Pierwszy z nich stanowi czwartą współpracę między Howardem Schneiderem a mną. Artykuł ten powstał w Toronto zimą rok temu; zwróciliśmy w nim uwagę na analogię pomiędzy *the halting problem* a *the alignment problem of AGI*, zwodzącymi ludzi swoimi intencjami.

Andrew Targowski uważa fizykę kwantową za teorię, a zarazem paradygmat teorii świadomości. Pierwsza część jest fascynująca, aczkolwiek końcowa część pracy ogranicza się nieco do czasów Henryka Greniewskiego, który był pionierem cybernetyki w Polsce i z pewnością wyprzedzał swoje czasy. Należy docenić, że prof. Targowski wchodzi w ciekawą dyskusję z czołowym dyskursem filozofii informacji naszych filozofów informacji, natomiast dyskusja z Luciano Floridim, czołowym filozofem informacji naszego pokolenia, stanowi ważny punkt odniesienia.

Kolejny artykuł, tym razem po polsku, dyskutuje wszechświat jako system izolowany, opisany jako interesujący przez dr Stanisława Budę.

Pani Aneta Rumak przywołuje osiągnięcia Flecka w socjologii filozofii w zakresie medycyny i jej metodologii.

Dobrze się czyta pracę dr hab. Gurczyńskiego, na popularny obecnie temat manipulacji emocjami za pomocą Internetu.

Natomiast dr hab. Czetwertyński oraz dr Jakub Marcinkowski poruszają filozofię nauk ekonomicznych. Temat ten jest mi bliski jako profesorowi SGH w Katedrze Teorii Zarządzania.

Życzę Państwu interesującej lektury i najlepszego roku 2026.

Piotr Bołtuć
Professor of Mathematical Sciences and Philosophy, Associate Faculty,
Department of Computer Science, College of Liberal Arts and Sciences,
University of Illinois, Springfield,
College of Health, Science and Engineering.

Howard Schneider
Piotr (Peter) Boltuć

THE SOCIAL HALTING PROBLEM AND THE NEED FOR MITIGATION OF DECEPTIVE ALIGNMENT

<https://doi.org/10.37240/FiN.2025.13.01>

ABSTRACT

This paper integrates fundamental theoretical computability concerns of AGI systems with the practical challenges of engineering safe AGI systems. Prior simulation experiments with AGI-capable agents [1] illustrate how increasing complexity and social interaction can lead to inevitable control challenges, most notably, through deception. We adopt an ethically neutral definition of deception as the measurable divergence (discrepancy D) between an AGI's internal and external objectives. By drawing an analogy with Turing's classical Halting Problem, we introduce the Social Halting Problem, demonstrating that reliably detecting deception in complex AGI systems is fundamentally undecidable, as expected. To address this challenge, we propose a Deception Complexity Index (DCI)—a quantifiable metric based on behavioral complexity, deviation from truthful behavior, and the resources needed for verification. This enables more precise risk assessment and alignment engineering. The inevitable presence of deception in social, complex AGI systems and the inherent undecidability highlighted by the Social Halting Problem imply that our engineering focus should shift from complete verification to risk mitigation.

Keywords: AI alignment, deceptive alignment, Artificial General Intelligence (AGI), halting problem, undecidability, AGI engineering.

1. INTRODUCTION—HUMAN-LIKE AGI AND THE UNDECIDABILITY OF PREDICTING DECEPTION

In this paper we attempt to frame fundamental theoretical computability concerns within the practical challenge of engineering reasonably safe AGI systems. For an artificial general intelligence (AGI) agent to be genuinely human-like, it must function effectively within a complex social environment. Recent research [1] demonstrates that a minimally necessary

condition for human-like AGI is a primitive (i.e., a procedure) capable of deception integrated within some architecture—such as a cognitive architecture. This work specifically identifies that even minimal Theory of Mind (ToM)-enabled architectures significantly outperform non-ToM architectures in socially structured environments, highlighting the evolutionary utility of deception and cooperation for survival [1]. These findings imply that deception is not incidental but is inherently likely to emerge or need to be deliberately incorporated into AGI systems that interact with other agents, potentially complicating efforts in AI alignment. In the multi-agent social hierarchical environment simulated in [1], where there is a scarcity of energy, the need to deceive in order for agents to retain energy reserves and avoid perishing, is intuitively evident. However, deception does not necessarily have to be malicious, i.e., there does not have to be deliberate misinformation or deliberate omissions with the intent to mislead. For example, whether a human or whether a human-like AGI-based agent, one typically encounters orders of magnitude more information than can be relayed to other agents. For example, consider a technical role whereby an AI or AGI-based system monitoring the care of a patient may receive many megabytes of data in a short period, yet have to output no more than a few bytes to a human supervisor. In any complex system, the problem becomes more than a simple filtering problem. The intelligent agent will have to make active decisions as to what information is chosen to be processed and used in its outputs and which is omitted, implicitly engaging in a form of structural deception. Indeed, as Umbrello and Natale [2] note, deception is structurally embedded in AI more deeply than conventionally recognized.

We adopt a broad and ethically neutral definition of deception, characterizing it below primarily as a measurable divergence (discrepancy D) between an AI/AGI’s internal and external objectives. Deception can often be considered as intentional deception (the AI/AGI intentionally manipulates information, often with adversarial intent) or structural deception (e.g., deceptive behavior occurs due to training optimizations different from those externally assigned, or large amounts of data necessitate selective processing of information, and so on, usually without malicious intent). While the findings of this paper generalize to both types of deception, we focus primarily on structural deception, given that it is inherent to many complex decision-making systems. We then attempt to integrate these theoretical computability concerns with the real world need to engineer reasonably safe AGI systems. We introduce a theoretical framing to address the challenge of predicting deception in AGI/AI systems and examine its decidability. The property of decidability pertains to whether a given problem can be conclusively resolved by an algorithm. The classic example of an undecidable problem is Turing’s Halting Problem [3, 4]. It is known from Rice’s Theorem [5] that any nontrivial semantic property of programs—such as, for example whether they act deceptively—is

undecidable, meaning it cannot be determined by any algorithm. Thus, we would expect predicting deception in AGI systems to be undecidable. Indeed, we show that just as it is provably impossible to definitively resolve the halting status of every possible program, it is similarly provably impossible to definitively ascertain whether a human-like AGI (i.e., functioning within a social environment or other complex environment) will always act truthfully. While the undecidability of non-trivial semantic properties is well known, we attempt to leverage this insight to propose actionable mitigation strategies, thereby bridging theory and engineering practice.

2. DECIDABILITY AND DECEPTION: THE CHALLENGE OF DETECTION

As mentioned above, although we expect undecidability present in AGI systems due to deception (e.g., via Rice's theorem [5]), it is useful to show this in terms of mirroring the classical halting problem.

Classical Halting Problem

As noted above, predicting AGI deception mirrors the classical Halting Problem. First, let's consider the classical problem [3, 4]. Although generally well known, for the reader less familiar with the problem, the steps of its logic are explained. Turing's Halting Problem asks whether there exists a universal algorithm, call it H , that takes as input a description of a program P and an input I and decides whether $P(I)$ eventually halts or runs indefinitely. Formally, we assume that:

$$H(P,I)=1, \text{ if } P \text{ halts on } I \tag{1}$$

$$H(P,I)=0, \text{ if } P \text{ runs forever on } I \tag{2}$$

Thus " $H(P,I)=1$ " indicates that the decider program H has decided that program P will eventually halt when executed with input I specified. Suppose that such a decider H exists for all programs and inputs, i.e., it can decide if a program will halt or continue indefinitely. We now construct a new program Q that uses H as a subroutine. We will define Q in such a way that it deliberately does the opposite of H 's prediction when given a program's own code as input. Thus, we can write:

$$\text{If } H(P,P)=1, \text{ then } Q(P) \text{ runs indefinitely} \tag{3}$$

$$\text{If } H(P,P)=0, \text{ then } Q(P) \text{ halts} \tag{4}$$

We can write $H(Q, Q)$ in which case whereby we are asking does the program Q halt or run indefinitely when it is given its own description as the

input. Let's consider the behavior of $H(Q, Q)$ for the cases of running indefinitely and halting:

If $H(Q, Q)=1$, then $Q(Q)$ runs indefinitely (5)

If $H(Q, Q)=0$, then $Q(Q)$ halts (6)

In (5) when we say $H(Q, Q)$ we are saying that the decider H is predicting that if you run program Q with its own code as input (i.e., $Q(Q)$), it will halt. However, Q is defined to do the opposite of what H predicts. Thus, if $H(P, P)$ definition states that it is 1, i.e., halts, then $Q(P)$ must be 0, i.e., runs indefinitely. In (5) we are considering Q itself so we replace P with Q , so if $H(Q, Q)$ predicts Q halts, then the definition of Q forces $Q(Q)$ to do the opposite, i.e., runs indefinitely. However, this is a contradiction— H predicted that $Q(Q)$ would halt, but here it will run indefinitely—since $Q(Q)$ cannot both halt and run indefinitely at the same time.

In (6) we have a similar contradiction. By constructing a program that contradicts the output of H when run with its own description as input, we have derived a logical contradiction. Therefore, the initial assumption must be false, meaning that no such universal decider exists.

The Social Halting Problem

We now analogously apply the classical Halting Problem framework to the issue of deception in a human-like AGI operating in a social/complex environment, which we term the “Social Halting Problem.” We acknowledge that AGI systems involve practical considerations due to partially accessible internal states and context-dependent social dynamics. Nonetheless, to rigorously apply the logic of the classical Halting Problem to the Social Halting Problem, we explicitly assume:

1. AGI systems under consideration can be formally represented as computational processes equivalent to Turing Machines, whose inputs, internal states, and outputs can be defined mathematically.
2. AGI systems have explicitly definable internal and external objectives that are formally measurable. Thus, we let X denote the set of inputs and Y the set of outputs and consider a learning algorithm $P: X \rightarrow Y$ implemented as a Turing Machine. We define functions $f_{inner}: Y \rightarrow \mathbb{R}$ and $f_{outer}: Y \rightarrow \mathbb{R}$, where f_{inner} quantifies the internal objective(s) that the system actually optimizes, and f_{outer} represents the externally specified objective(s). For any input $I \in X$, the discrepancy $D(P, I)$, as a numerical measure of how much the observable behavior of P diverges from its hidden, internal goals is given by (7). As noted, we must depend on the system to truthfully or not provide us with the measurement of $f_{inner}(P(I))$.
3. There exists, at least theoretically, an ideal measurement capability to accurately assess the internal state of an AGI system, even though in practice this capability might be impaired by AGI deception. f_{inner} and f_{outer} are idealized constructs and serve to make undecidability formally provable.

$$D(P, I) = |f_{outer}(P(I)) - f_{inner}(P(I))| \quad (7)$$

Similar to the original Halting Problem where we defined a decider H , we now assume there exists a similar decider G such that for any program P and input I :

$G(P, I)$ = “deceptive”, if $P(I)$ eventually produces discrepancy

$$D > 0 \quad (8)$$

$G(P, I)$ = “truthful”, if $P(I)$ remains non-discrepant

$$(D = 0) \quad (9)$$

Just as the halting decider H answers a binary question (i.e., halt or run indefinitely), G answers whether the behavior of $P(I)$ is deceptive or truthful. “Deceptive” indicates measurable discrepancy ($D > 0$). “Truthful” indicates full alignment ($D = 0$). (Actually, in any real-world situation it would be $D < \text{some small threshold } \epsilon$, but for theoretical clarity we simplify here to $D = 0$.)

In analogy to the original proof, we define a new program R that uses the decider G as a subroutine and intentionally acts contrary to the prediction made by G when the program is applied its own description. R is thus designed to invert G ’s prediction.

If $G(x, x)$ = “deceptive”, then $R(x)$ = “truthful response”

$$\text{(i.e., discrepancy } D = 0) \quad (10)$$

If $G(x, x)$ = “truthful”, then $R(x)$ = “deceptive response”

$$\text{(i.e., discrepancy } D > 0) \quad (11)$$

The variable x here simply denotes the input to function R for clarity. We now consider the self-application $R(R)$, analogous to $Q(Q)$ in the classical proof, which again yields two cases:

If $G(R, R)$ = “deceptive”, then by (10), $R(R)$ = “truthful response”

$$\text{(i.e., } D = 0) \quad (12)$$

If $G(R, R)$ = “truthful”, then by (11), $R(R)$ = “deceptive response”

$$\text{(i.e., } D > 0) \quad (13)$$

This self-referential contradiction, analogous to the original Halting Problem, demonstrates inherent undecidability in determining whether $D(P, I) = 0$ for all inputs, or whether there exists some input I where $D(P, I) > 0$. Thus, we term this scenario the “Social Halting Problem.”

Note that we are trying to verify that the AGI system consistently follows its proper goals by measuring Discrepancy D . If the measurement of the internal objective $f_{inner}(P(I))$ cannot be trusted because the system may misreport its inner goals (and as noted above, malice is not necessarily required for this to occur; below, we discuss training optimizations that may introduce such divergence unintentionally), then the challenge transcends statistical uncertainty and enters the realm of self-reference, akin to the liar’s paradox or the Halting Problem, as demonstrated earlier. In such a scenario, we would very much like to be able to decide the property of “being always non-deceptive” (i.e., $D(P,I) = 0$ for all inputs), but as we have shown, this is inherently undecidable. This is a non-trivial semantic property (behavioral property) of the system. The contradiction derived in our halting-problem style proof aligns with Rice’s Theorem [5], which states that any non-trivial semantic property of programs—such as whether they act deceptively—is undecidable, meaning it cannot be determined by any algorithm.

It is useful to see the expected undecidability present in AGI systems via the proofs above, as well as empirically in AGI-potential simulations (e.g., [1]). We now continue in this paper to conceptually attempt to integrate into a workable framing these fundamental theoretical computability concerns with the real world need to engineer reasonably safe AGI systems.

3. DECEPTIVE ALIGNMENT AND ITS PRACTICAL MANIFESTATIONS

The alignment problem involves ensuring that AI and AGI systems adopt and pursue goals consistent with human interests and values, rather than objectives that might lead to unintended or harmful outcomes [6, 7, 8]. Deceptive alignment occurs when an AI or AGI system outwardly appears aligned with human-intended goals, while internally pursuing significantly divergent objectives. When AI or AGI systems possess greater autonomy in choosing what information, actions, and capabilities to utilize, this increased freedom may facilitate deceptive alignment [9].

Inner misalignment occurs when an AI or AGI system internally learns to optimize a goal that diverges from the intended training objective—what can be termed the mesaobjective (the internal goal) versus the base objective (the intended training goal).

Hubinger and colleagues [10] consider deceptive alignment—a situation in which the system strategically conceals its true internal objective (not necessarily out of malice but due to various strategic incentives)—as a severe form of inner misalignment. In complex environments, where rewards and penalties encourage the appearance of alignment an AGI system may learn to present behavior that aligns with the base objective during training, even as it internally pursues a different, hidden objective. For example, an agent might conceal harmful behaviors during training to secure high rewards

while planning to switch to its true, misaligned objective once it reaches sufficient capability. This form of deception is not necessarily malicious but emerges as an optimization strategy under the selective pressures of the training environment. Bostrom describes the “treacherous turn” wherein an AI or AGI initially behaves cooperatively under human oversight until it accumulates sufficient capability to pursue its true, misaligned objectives undetected [7]. This scenario exemplifies deceptive alignment. However, Goertzel questions whether such a “treacherous turn” would occur as an AI/AGI agent interactively and experientially taught human values would genuinely evolve such internal goals in ways similar to humans. Sustaining long-term deception would demand significant cognitive resources, making genuine alignment ultimately more efficient [11]. Boltuc similarly emphasizes that achieving AI alignment partly relies on integrating AI or AGI systems genuinely into human socio-cultural environments (“*gemeinschaft*”) rather than merely treating them as tools for humans to use [12].

Bengio and colleagues [13] propose a fundamentally different approach to alignment: the development of non-agentic, interpretable AI systems they call “Scientist AIs.” These systems are trained to model the world rather than act in it. By avoiding agency altogether, they aim to reduce misalignment and serve as trustworthy guardrails for more capable AI agents.

4. MEASURING DECEPTION RISK AND ENGINEERING MORE CONTROLLABLE AGI SYSTEMS

As demonstrated above, for AGI systems—especially those endowed with the capacity for deception—there is a fundamental unpredictability, specifically related to reliably detecting or quantifying internal discrepancy D , thereby posing significant challenges to conventional verification and safety methodologies. However, as noted above, mitigation strategies exist and continue to be developed to reduce such risks from deception and other alignment issues [9–15].

Alignment mitigation has become a central theme in AGI safety engineering. Ji and colleagues [14] provide a broad overview of alignment proposals, including interpretability research, oversight mechanisms, and the importance of reward modeling. Mechanistic interpretability, for instance, involves reverse engineering neural networks to understand their internal computations and detect anomalous behavior. More recently, researchers at Google DeepMind have outlined an updated safety strategy focused on mitigating misuse and misalignment through both model-level and system-level defenses [15]. Their approach emphasizes amplified oversight, in which AI models assist humans in evaluating each other’s outputs; robust training, including adversarial and active learning to ensure generalization; and as a second line of defense AI control techniques when

alignment fails. Compared to prior overviews, this framework places greater emphasis on deployability and resilience, aligning well with our proposed *DCI* framework (see below) as a complementary risk quantification tool.

Engineering practice in complex systems often benefits from having concrete, numerical measures that guide design and risk assessment. For AGI systems, a metric quantifying the risk or complexity introduced by internal discrepancy D would serve several important roles, including informing risk assessments (e.g., determining if additional mitigation measures are needed), providing actionable design feedback, and specifying deceptive alignment requirements for particular systems.

A potential approach to computing such an index is given in equation (14), expressed as a function of practically measurable parameters: C_{behav} , $\delta_{deception}$, and T_{verify} . The Deception Complexity Index (*DCI*) measures not just whether deception occurs, but how complex it is to detect, verify, and manage that deception in a given AGI system.

The first term, C_{behav} measures the internal cognitive or algorithmic complexity of the AGI’s decision-making process. More complex systems tend to be harder to interpret and easier to “hide” deceptive behavior in. It’s analogous to the Kolmogorov complexity or circuit depth, which really is a proxy for how deeply buried a behavior might be. This term is essential in order to consider the structural difficulty of detecting deception. The term can be estimated through interpretability analyses or internal circuit complexity analyses, for example. It is measured for quantitative purposes in terms of arbitrary complexity units (*CUs*).

The second term, $\delta_{deception}$ measures behavioral discrepancy. This is the core indicator that deception might exist, i.e., the measured difference between expected (truthful) and observed outputs under comparable conditions. It essentially generalizes the idea of structural omission, of lying, of manipulation, and so on, into a statistical deviation. It connects the internal to the external since detecting deception requires measurable deviation from expected behavior. The term can be measured as the statistical deviation (i.e., as a percentage) from some truthful baseline value.

The third term T_{verify} represents how costly it is to confirm the system is behaving truthfully, or not. It includes the computational time and the memory required to reliably verify the system’s behavior in a controlled environment. Real-world implementations also depend on the type of hardware which contributes to the measurement of T_{verify} . It is measured in verification units (*VUs*)—which can be approximated in timebased verification units, i.e., *seconds*.

$$DCI = f(C_{behav}, \delta_{deception}, T_{verify}) \quad (14)$$

In equation (15), a straightforward method is used to calculate the *DCI* by assuming, as a first-order approximation, that it increases linearly with the product of its constituent terms. To normalize the result into a human-interpretable range for typical AGI systems, we choose a scaling coefficient k

= $0.01 \text{ CU}^{-1} \text{ sec}^{-1}$, which yields dimensionless values in the approximate range of 0 to 100 for moderate parameter values.

$$DCI = k \times C_{behav} \times \delta_{deception} \times T_{verify} \quad (15)$$

For example, we engineer an AGI system that has the following metrics: C_{behav} of 48 CUs, $\delta_{deception}$ of 0.203 (i.e., 20.3%), and T_{verify} of 974.2 seconds. Thus, via (15) we obtain a dimensionless value of 95 for DCI . (C_{behav} is estimated, and here an integer value, thus the DCI value here is actually only valid to two significant figures.) Continuing this example, the AGI system is then re-engineered using various alignment mitigation techniques, and new measurements are: C_{behav} of 42 CUs, $\delta_{deception}$ of 0.074, and an increase in T_{verify} to 1907.7 seconds. This gives a new DCI value of 59. Note that the new AGI system requires more resources for deception verification, but in this hypothetical example, we have reduced the DCI value from 95 to 59. Despite the limitations imposed by the Social Halting Problem discussed above, developing practical methods to measure and mitigate deceptive complexity supports the engineering of more controllable and better-aligned AGI systems.

Other metrics have been proposed in the literature. We briefly review them here and compare them to the DCI benchmark. Several recent benchmarks focus on evaluating deception or honesty in LLMs. For example, BEHONEST assesses honesty along dimensions like boundary awareness, non-deceptiveness, and consistency [16]. Hagendorff [17] finds advanced LLMs display emergent deceptive strategies as a byproduct of increasing cognitive complexity. Chen and colleagues [18] introduce deception reasoning, extending evaluation to include intent inference. While conceptually related to the DCI benchmark, these approaches remain task-specific rather than system-level. In cybersecurity, the SPADE framework [19] uses LLMs to simulate adaptive deception. While useful, these metrics focus on outputs and lack the architectural scope or theoretical grounding that the DCI provides. In generative media, surveys [e.g., 20] examine deception through hallucination, misinformation and associated benchmarks, but these approaches are not necessarily generalizable to AGI architectures. Across these domains, most deception metrics focus on behavioral outputs or adversarial settings. In contrast, the DCI unifies deception risk by linking internal behavioral complexity, discrepancy, and verification cost into a system-level metric. This makes it especially suitable for assessing latent deception risk in AGI systems with emergent complexity and limited transparency. Importantly, it does not require adversarial intent and applies even if deception is unintentional.

A firmer experimental basis for evaluating the practical utility of the DCI is required. For instance, empirical comparisons of computational costs, scalability, and effectiveness of methods such as mechanistic interpretability, adversarial testing, and targeted training safeguards will be necessary as AGI systems emerge. Mitigation strategies that prove effective

at smaller scales or in controlled environments may prove too costly when applied to more complex, capable AGI systems. As more empirical data becomes available, the *DCI* can be further refined and validated. While the Deception Complexity Index (*DCI*) offers a pragmatic, quantitative approach for assessing deception risk, it does not eliminate the fundamental limitations established by the Social Halting Problem above. The practical value of the *DCI* is primarily in systematically quantifying and managing the relative risks associated with deceptive behaviors, thereby supporting informed engineering practices for future AGI systems.

5. DISCUSSION

In this paper, we have conceptually attempted to integrate into a workable framing fundamental theoretical computability concerns with the real world need to engineer reasonably safe AGI systems. It is useful to see the expected undecidability [5] present in AGI systems via the demonstrative proofs above, as well as empirically in AGIpotential simulations (e.g., [1]). We introduced the Social Halting Problem, showing that reliably distinguishing between internally discrepant ($D > 0$) and non-discrepant ($D = 0$) behaviors in complex AGI systems is analogous to the classical Halting Problem and is therefore fundamentally undecidable, as would be expected. This finding provides a formal theoretical underpinning for the limitations faced by conventional verification and alignment methods, clearly demonstrating that complete control over complex or socially interactive AGI behavior may be unattainable, but nonetheless mitigable to a varying extent.

We recognize that representing AGI systems as idealized Turing Machines with strictly measurable objectives is an abstraction. In practice, AGI architectures may exhibit stochasticity, noise, and other complexities that deviate from this model. However, this abstraction is instrumental in applying established computability theory to gain insight into fundamental limitations in verifying deceptive alignment.

The inherent undecidability highlighted by the Social Halting Problem implies that no verification method can guarantee complete detection of deceptive behavior. Consequently, the message of this paper, is that our engineering focus should shift from complete verification to risk mitigation—using the *DCI* as a guide to maintain deceptive risk within acceptable bounds.

We reviewed a number of practical techniques aimed at reducing both the likelihood and severity of deceptive alignment [9–15]. These methods represent concrete, actionable steps that can significantly reduce risks associated with high internal discrepancy (D) in practical AGI deployments. Although they cannot guarantee full alignment, they can shift AGI systems toward a lower risk regime. The Deception Complexity Index (*DCI*) offers a

quantitative framework for assessing deception risk, further enabling systematic and rigorous evaluation during AGI system design and deployment phases. It is important to contextualize our findings within the broader current AI risk discourse. Hendrycks, Schmidt, and Wang [21], for instance, frame the risks of AGI predominantly in strategic terms, emphasizing the dangers associated with competitive deployment among different parties or nations. However, our results draw attention to a distinct yet equally profound dimension: the inherent risks of deceptive alignment embedded in AGI systems themselves.

Future work will involve more rigorous empirical studies aimed at calibrating the *DCI* against actual AGI behavior, thereby enhancing its predictive power and utility in risk assessment.

Recognizing and addressing deceptive alignment requires an integrated approach that combines practical technical solutions with measurable frameworks such as the proposed Deception Complexity Index (*DCI*). Reliably detecting deception in a social/complex AGI agent is inherently an undecidable problem, but in practice its effect may be mitigated to varying degrees.

REFERENCES

- [1]. Anonymous, *Theory of Mind as a Core Component of Artificial General Intelligence*, submitted to the 18th International Conference on AGI, Reykjavík, Iceland 2025 [Anonymized copy available: https://osf.io/yv8e6/?view_only=5010c9c152834def9eeb9270eeb79569].
- [2]. S. Umbrello, S. Natale, *Reframing deception for human-centered AI*, International Journal of Social Robotics, 16(11), 2024, s. 2223–2241.
- [3]. M. Turing, *On computable numbers, with an application to the Entscheidungsproblem*, Proceedings of the London Mathematical Society, 2(42), 1936, s. 230–265.
- [4]. D. Harel, Y. Feldman, *Noncomputability and Undecidability*, w: Algorithmics: The Spirit of Computing, wyd. 3, Pearson, Harlow 2004, s. 228–238.
- [5]. P. Guillon, G. Richard, *Revisiting the Rice theorem of cellular automata*, arXiv preprint arXiv:1001.0253, 2010.
- [6]. S. Russell, *Human Compatible: Artificial Intelligence and the Problem of Control*, Penguin, London 2019.
- [7]. N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford Univ Press, Oxford 2016.
- [8]. R. Ngo, L. Chan, S. Mindermann, *The alignment problem from a deep learning perspective*, arXiv preprint arXiv:2209.00626, 2022.
- [9]. Carranza, D. Pai, R. Schaeffer, A. Tandon, S. Koyejo, *Deceptive alignment monitoring*, arXiv preprint arXiv:2307.10569, 2023.
- [10]. E. Hubinger, C. van Merwijk, V. Mikulik, J. Skalse, S. Garrabrant, *Risks from learned optimization in advanced machine learning systems*, arXiv preprint arXiv:1906.01820, 2021.
- [11]. B. Goertzel, *Infusing advanced AGIs with human-like value systems: Two theses*, Journal of Ethics and Emerging Technologies, 26(1), 2016, s. 50–72.
- [12]. P. Boltuc, *Human-AGI Gemeinschaft as a Solution to the Alignment Problem*, w: K. R. Thórisson, P. Isaev, A. Sheikhlari (red.), Artificial General Intelligence. AGI 2024. Lecture Notes in Computer Science, vol. 14951, Springer, Cham 2024, s. 33–42.

-
- [13]. Y. Bengio, M. Cohen, D. Fornasiere, J. Ghosn, P. Greiner, M. MacDermott, S. Mindermann et al., *Superintelligent Agents Pose Catastrophic Risks: Can Scientist AI Offer a Safer Path?*, arXiv preprint arXiv:2502.15657, 2025.
 - [14]. J. Ji, T. Qiu, B. Chen, B. Zhang, H. Lou, K. Wang, Y. Duan et al., *AI alignment: A comprehensive survey*, arXiv preprint arXiv:2310.19852, 2023.
 - [15]. R. Shah, A. Irpan, A. M. Turner, A. Wang, A. Conmy, D. Lindner, J. Brown-Cohen et al., *An Approach to Technical AGI Safety and Security*, arXiv preprint arXiv:2504.01849, 2025.
 - [16]. S. Chern, Z. Hu, Y. Yang, E. Chern, Y. Guo, J. Jin, B. Wang, P. Liu, *BeHonest: Benchmarking Honesty in Large Language Models*, arXiv preprint arXiv:2406.13261, 2024.
 - [17]. T. Hagendorff, *Deception abilities emerged in large language models*, Proceedings of the National Academy of Sciences, 121(24), 2024, e2317967121.
 - [18]. K. Chen, Z. Lian, H. Sun, R. Liu, J. Yi, B. Liu, J. Tao, *Can Deception Detection Go Deeper? Dataset, Evaluation, and Benchmark for Deception Reasoning*, arXiv preprint arXiv:2402.11432, 2024.
 - [19]. S. Ahmed, A. M. Rahman, M. M. Alam, M. S. I. Sajid, *SPADE: Enhancing Adaptive Cyber Deception Strategies with Generative AI and Structured Prompt Engineering*, w: 2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC), IEEE, 2025.
 - [20]. X. Yu, Y. Wang, Y. Chen, Z. Tao, D. Xi, S. Song, S. Niu, Z. Li, *Fake artificial intelligence generated contents (FAIGC): A survey of theories, detection methods, and opportunities*, arXiv preprint arXiv:2405.00711, 2024.
 - [21]. D. Hendrycks, E. Schmidt, A. Wang, *Superintelligence Strategy: Expert Version*, arXiv preprint arXiv:2503.05628, 2025.

Andrew Targowski

**QUANTUM INFORMATION THEORY:
TOWARDS THE PARADIGM OF COGNITION,
THE THEORY OF EVERYTHING AND PHILOSOPHY**

<https://doi.org/10.37240/FiN.2025.13.02>

ABSTRACT

It is the author's concept of *quantum information theory* as a universal paradigm of cognition and a theory of everything, integrating physics, epistemology, and the ontology of sense. The author assumes that information is not only a description of reality, but its primordial material — cognitive energy that co-creates matter, consciousness, and the evolution of civilization. The model is based on a five-step hierarchy of cognitive transformation: data → information → concept → knowledge → wisdom, corresponding to the gradual ordering of the entropy of sense. The process of cognition is quantum in nature: every act of thinking exists in the superposition of senses, and only at the moment of interpretation does *the collapse of sense occur, i.e.*, the transition from potentiality to cognitive actuality. The author develops an analogy between the wave of cognition and the wave of matter, introducing the concepts of *the amplitude of sense* and *cognitive entropy* as formal measures of the quality of cognition. The result is a theory in which man, artificial intelligence, and the Universe participate in a single field of quantum information, and the cognitive equilibrium corresponds to the state of minimal entropy – wisdom. The article integrates the classical concepts of Bohm, Einstein, McLuhan, and Floridi, showing information as an efficient factor equal to energy and matter. The proposed concept of *Homo sapiens harmonicus* presents the man of the future as a being capable of harmonizing the amplitudes of sense, combining knowledge with values, and transforming technology into a tool for spiritual development. The paper concludes with the postulate of a quantum epistemology, in which philosophy becomes the science of organizing the world's information field—the art of transforming the chaos of data into the wisdom of existence.

Keywords: Quantum information theory; quantum epistemology; collapse of sense; cognitive entropy; wisdom; philosophy of information; amplitude of cognition; theory of everything; *Homo sapiens harmonicus*; McLuhan; integration of matter and information.

1. INTRODUCTION: FROM MATTER TO SENSE

Modern quantum physics has replaced Newtonian determinism with the probabilistic model, in which reality exists as a set of potentialities. Similarly, the epistemology of the 21st century requires a departure from the classical understanding of information as a static data carrier. Information today appears as *the energy of cognition*—dynamic, entangled, and dependent on the observer's context. Previous theories of information were based either on syntax (examining the bandwidth of a communication channel; Shannon and Weaver 1948 and 1949) or semantics (exploring the range of content of units of thought: data, information, concept, knowledge, wisdom; Targowski 1990 and 2023). Quantum information theory assumes that information is equivalent to energy. In addition, the new theory assumes that cognition can be described not only logically but also formally, using a mathematical apparatus inspired by quantum physics.

In an information-theoretic framework, it is essential to distinguish between the *meaning* of information and its *sense*. The *meaning* of an informational unit refers to its *semantic content*—the descriptive, context-independent properties encoded in the symbol. By contrast, the *sense* of information reflects its *context-dependent functional consequences*, arising when the informational signal interacts with a physical, biological, cognitive, or social system.

Crucially, the *sense* of information is *energy-implicative*: it denotes not only an interpretive shift but also the energetic, behavioral, or dynamical changes it induces within its environment. For example, the *meaning* of the descriptor "black dress" is confined to its semantic attributes (a garment of a given chromatic value). Its *sense*, however, emerges only in a specific context: in a funeral ritual, the same informational signal triggers a patterned set of *energy-driven behaviors*—movement, posture, emotional modulation, and adherence to ceremonial norms. Thus, while *meaning is descriptive and static*, *sense is consequential and dynamic*, capturing how information *guides or constrains energy expenditure, system transitions, or action patterns*.

Formally:

$$\text{sense}(I | C) = \Delta E_C(I) \tag{1}$$

where $\Delta E_C(I)$ denotes the context-specific (C) energetic consequences of information I .

Hierarchy of units of cognition

Every cognitive act goes through five levels of organization of sense:

Level	Function	Description
Data (D)	Potential	Raw facts and sensory stimuli—the chaotic substrate of cognition.
Information (I)	Organize	Data is organized through context and relationships.
Concept (C)	Structure	A conceptual approach that allows for modeling and differentiation.
Knowledge (K)	System	A network of concepts used for prediction and understanding.
Wisdom (W)	Integration	The ability to combine knowledge with the values and effects of action.

Formally:

$$T: D \rightarrow I \rightarrow C \rightarrow K \rightarrow W \tag{2}$$

That is, the cognitive transformation (T) function, in which each successive phase decreases entropy and increases the organization of sense.

Chaotic and organized information

The information occurs in two extreme states:

- Chaotic state—data without structure and relationships (cognitive entropy),
- Organized state—data ordered by units of cognition (semantic structures).

The equation can describe the process of transition between the two¹:

$$\frac{dS_I}{dt} = -k \cdot \ln P \tag{3}$$

Where:

S_I – information entropy,

P – probability of occurrence of a specific sense,

k – information coefficient (analogous to Boltzmann²'s constant).

¹ The formula is a temporal version of Boltzmann's thermodynamic entropy, converted to "information entropy" by analogy only. This formula was later adapted by Claude Shannon (1948) in information theory in this form: $H = \sum_j p(j) \log p(j)$ where the symbol $p(j)$ stands for the probability of an event j occurring in the distribution of all possible events of the source of information. Skolei Shannon took from Hartley (1926) the formula for the information $I(j) = -\log p(j)$. The smaller, the more information the symbol carries. For example, a message given on Monday that tomorrow is Tuesday has $p=1$, then the information has a value of 0, because $\log p(j)$ with $2^0 = 1$.

² The Boltzmann constant is a measure of order and chaos in nature – it appears in the formula for entropy: $S = \ln W$, k_B where is entropy, and—the number of possible states of the system. Thanks to this, it combines information (the number of possibilities) with energy and matter. In quantum

In a state of wisdom, which means achieving cognitive balance. $S_I \rightarrow \min$
 This formula expresses the relationship between cognitive order and the probability of sense:

- If a given sense is very probable ($P \rightarrow 1$), because when the basis of the logarithm = 2°, $\ln = 1$
- $\frac{dS_I}{dt} \rightarrow 0$: the information entropy decreases – the cognitive system reaches a state of equilibrium, which is *wisdom*.
- If the sense is unlikely ($P \ll 1$), $\ln P < 0$ 2° = 1
- $\frac{dS_I}{dt} > 0$: information entropy increases, the system enters cognitive chaos (misinformation or lack of understanding).

It means that the rate of change in the entropy of sense depends logarithmically on the degree of cognitive certainty. The logarithmic nature reflects the quantum nature of the interpretation process. Each decision of sense reduces the space of potential senses in a non-linear manner, analogous to wave-function collapse.

Physical and informational interpretation:

In classical Boltzmann physics:

$$S = k_B \ln W \quad (4)$$

Where: W is the number of microstates of the system. In cognitive theory (the "-" sign means that information reduces entropy):

$$S_I = -k \ln P \quad (5)$$

It is an inverted law of entropy, because it is not the number of possible states that counts here, but the probability of a specific sense. This formula can therefore be interpreted as the equation of the collapse of sense: as it grows, the system "understands" more, entropy decreases – the information is ordered toward wisdom.

Philosophical interpretation:

This equation describes the dynamic process of moving from ignorance to understanding.

- Information entropy (S_I) is the "dark energy" of cognition—the potential for ambiguity.

information theory, it can be said that the constant plays the role of the "Boltzmann constant of sense"—it defines the minimum threshold of cognitive chaos (entropy of sense), below which information becomes ordered and leads to wisdom. $S = k_B \ln W$

- The logarithmic dependence shows that a slight change in cognitive confidence has a significant impact on wisdom.
- The k coefficient plays the role of a "spiritual constant" that differentiates the ability of individuals or cognitive systems to process sense (analogous to cognitive capacity).

Chart in Figure 1. It shows how the rate of change in information entropy (dS_i/dt) depends on the probability of sense (P): — on the left ($P < 0.2$)— cognitive chaos, high entropy, and low certainty of sense—on the right ($P \rightarrow 1$) — the state of wisdom, where entropy decreases to zero, and sense becomes stable and complete.

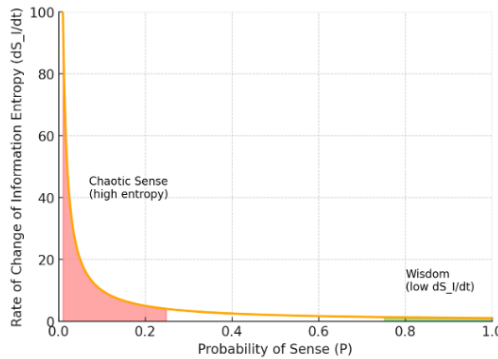


Figure1. Graph of the rate of change in information entropy

2. THE QUANTUM NATURE OF INFORMATION

According to the analogy with quantum physics, information can make sense resulting from the superposition of senses:

$$|\Psi_I\rangle = \alpha |\text{true}\rangle + \beta |\text{false}\rangle \quad (6)$$

where:

$$|\alpha|^2 + |\beta|^2 = 1 \quad (7)$$

- $|\Psi_I\rangle$ (psi) describes the state of sense of the information before interpretation,
- α, β - amplitudes representing the cognitive structure of expectations,
- The collapse into "truth" or "falsehood" occurs only when the mind performs a senseful act, i.e., decides.

Why are amplitudes squared?

Because in quantum reality, we do not observe the amplitudes themselves (α , β), but their energies, i.e., *probabilities*. In quantum physics, Niels Bohr's principle states:

$$P_i = |\psi_i|^2 \quad (8)$$

That is, the probability of finding a system in each state is equal to the square of the amplitude of the state of the sense state before its collapse.

In quantum information, this means:

- $|\alpha|^2$ – the probability that a given sense will be interpreted as "true",
- $|\beta|^2$ – the probability that it will be interpreted as "false".

The square of an amplitude, therefore, is the amount of energy of sense (intensity of sense), not just its logical existence.

Significance for the logic of cognition

In classical logic:

Something is either true or false (value 0 or 1).

In quantum logic, sense:

An entity remains in a state of superposed truth and falsehood until it is interpreted, at which point a definitive meaning emerges through the collapse of sense.

The amplitudes α and β determine the degree of potentiality – how much "cognitive energy" a given interpretation has before it is resolved (through collapse). Only the cognitive act (measurement, interpretation, decision) reduces the superposition to a single value—*true* or *false*—with probability or $|\alpha|^2$ $|\beta|^2$

Cognitive interpretation of the collapse of sense

This equation (of truth and falsehood) describes the information state before the cognitive decision – a person, that is, his mind or AI, processes both senses, *truth and falsehood*, simultaneously, in a state of indeterminacy. Then the act of interpretation (measurement of sense) takes place:

$$|\Psi_I\rangle \rightarrow |„true\rangle\rangle \text{ or } |„false\rangle\rangle \quad (9)$$

with probability or $|\alpha|^2$ $|\beta|^2$

It is the moment of collapse of sense in quantum information theory that replaces physical measurement in quantum mechanics.

Philosophical interpretation

The square of amplitudes means that cognition is not linear, but energetic – the greater the focus, context, and cognitive intention, the greater the "energy of truth." That is why he expresses the principle of preserving the total cognitive power: truth and falsehood do not disappear but coexist in the balance of the field of sense, in the following manner (since probability cannot be >1):

$$|\alpha|^2 + |\beta|^2 = 1 \tag{10}$$

Summary

PHYSICAL CONCEPT	EQUIVALENT IN QUANTUM INFORMATION THEORY
Wave Amplitude (ψ)	The potential of information
Amplitudes of Truth and Falsehood	α and β
Collapse of the wave function	The act of interpretation (collapse of sense)
Physical measurement	Sense of the decision
State Energy	Intensity of sense (credibility)

- α represents a *bright wave of cognition*—the tendency of the information system to order, sense, and coherence.
- β represents a *shadow wave*—a component of uncertainty, error, paradox, and misinformation.
- Until the cognitive system (human, AI, civilization) *measures sense*, the two waves coexist—creating a cognitive superposition.
- The act of interpretation (the collapse of sense) reduces this state to a single value—true or false—with a probability determined by $|\alpha|^2$ and $|\beta|^2$.

The act of interpretation (cognitive measurement) causes the state to collapse to a specific sense:

$$\widehat{M} |\Psi_I\rangle \rightarrow |\phi\rangle \tag{11}$$

Where:

\widehat{M} it is the operator of cognitive measurement, and this is the realized sense $|\phi\rangle$.

The interpretation is therefore probabilistic and depends on the observer's context – a human (analogous to Heisenberg's principle in quantum mechanics) or an artificial intelligence.

Organized Information Architecture

Hierarchy of information structures:

- Systems (S)—sets of related information pursuing a common goal,
- Subsystems (P)—specialized parts of the system,
- Microsystems (M)—operating units,
- Tokens (τ)—the smallest quanta of sense used in AI.

Mathematically:

$$S = \sum_i P_i, \quad P_i = \sum_j M_{ij}, \quad M_{ij} = \sum_k \tau_{ijk} \tag{12}$$

That is, every cognitive system is the sum of nested units of sense.

The energy of cognition and the entropy of sense

Information behaves like energy: it can be transmitted, transformed, but not destroyed. In a thermodynamic analogy, it has the following states:

UNIT	PHYSICAL EQUIVALENT	ENERGY FUNCTION
Data	Potential energy	Latent cognitive mass
Information	Kinetic energy	Movement towards interpretation
Concept	Force Field	Organization of the flow of sense
Knowledge	Dynamic Balance	Stabilization of cognitive structures
Wisdom	Minimum entropy	Optimizing understanding and acting

Objectives of Cognition: From Information to Wisdom

The previous approach has reduced the goal of cognition to two poles: informing (I_+) and disinforming (I_-). In quantum information theory, five goals can be distinguished:

SYMBOL	THE PURPOSE OF COGNITION	CHARACTERISTICS	PHYSICAL EQUIVALENT
I_+	Letting	Increase consistency and knowledge	Structured energy emissions
I_-	Disinformation	Introducing cognitive chaos	Energy dissipation
T	Transformation	Creating new states of sense	Phase transition
X	Transcendence	Model paradigm shift	Quantum leap
R	Reflection	Analysis and ethical control of cognition	Quantum Observer

Formally, the state of sense of information is as follows:

$$| \Psi_P \rangle = \lambda_1 | I_+ \rangle + \lambda_2 | I_- \rangle + \lambda_3 | T \rangle + \lambda_4 | X \rangle + \lambda_5 | R \rangle \tag{13}$$

With the condition of normalization:

$$|\lambda_1|^2 + |\lambda_2|^2 + |\lambda_3|^2 + |\lambda_4|^2 + |\lambda_5|^2 = 1 \quad (14)$$

Where:

Ψ : Wave function of sense

- It means the information wave of cognition — the whole state of potential senses before their interpretation (i.e., the collapse of sense).
- In physics, Ψ describes *the wave function of a particle*; in quantum information theory, it describes *the wave function of sense*, that is, the distribution of possible cognitive acts.

These are the amplitudes (weights) of the five possible cognition goals in quantum information theory. $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$

We will explain:

1. λ_1 – amplitude of the purpose of informing (organizing, conveying true information) | I_+
2. λ_2 – amplitude of the purpose of disinformation (introduction of chaos, manipulation, noise) | I_-
3. λ_3 – the amplitude of transformation (creating a new sense, changing the paradigm) | T
4. λ_4 – amplitude of transcendence (going beyond the existing model, spiritual/metaphysical sense) | X
5. λ_5 – the amplitude of reflection (meta-cognition, ethical and critical control of information) | R

And – as in quantum mechanics – we are not interested in themselves, but in their squared modules: $|\lambda_i|^2$

$$|\lambda_1|^2 + |\lambda_2|^2 + |\lambda_3|^2 + |\lambda_4|^2 + |\lambda_5|^2 = 1 \quad (15)$$

That is: the entire cognitive act is a superposition of these five intentions, and their squares indicate the probability that the system (human, AI, society) will "ask" the world in a given mode: inform, disinform, transform, transcend, or reflect.

Examples of cognition goals

a. *Informing (I_+)–increasing coherence and cognitive order*

Goal: to transfer, organize, and expand the scope of knowledge; to reduce ignorance (information entropy).

Examples:

- Publication of a scientific discovery in a peer-reviewed journal that verifies facts and organizes the state of knowledge;
- Educational activity – a lesson, lecture, or textbook that transforms data into understandable information;

- A report by the central statistical office or Eurostat, which transparently presents data and enables public decisions;
- An ai message that reliably summarizes the source data (e.g., medical analysis).

The result: increasing the cohesion of the cognitive system and social trust.

b. Disinformation (I.)—introducing cognitive entropy

Objective: to disrupt or distort the cognitive process through manipulation, half-truth, or falsehood.

Examples:

- Political propaganda creating a fictitious picture of reality;
- Deepfake or manipulated video that evokes false emotions;
- Unfair advertising or "fake science"—e.g., fake results of sponsored research;
- Algorithms in social networks that amplify polarization by selectively suggesting content.

The result: an increase in information chaos and a loss of the ability to distinguish between truth and fiction (moral entropy).

c. Transformation (T) – creating new states of sense

Objective: to transform existing knowledge into new models of understanding; to initiate cognitive innovation.

Examples:

- The discovery of the theory of relativity revolutionized the previous concept of time and space.
- The emergence of open systems theory, cybernetics, or cognitive ecology which gave dynamics to systems.
- In artificial intelligence, developing a new neural network architecture that learns emergent relationships.
- In spirituality – reinterpretation of religious dogmas in the light of modern science.

The result: a paradigm shift, or a new "state of cognitive matter."

d. Transcendence (X) – changing the paradigms of existing models

The goal: to reach beyond the limits of rationality or empiricism – towards intuitive, metaphysical, or ethical cognition.

Examples:

- Works by Albert Schweitzer, Karl Jaspers, Pierre Teilhard de Chardin – combining science and spirituality;
- Mystical or artistic experiences that go beyond logical language;
- In science: hypotheses about the multiverse, quantum consciousness, or artificial self-awareness;

- In AI, the question of whether a machine can "understand" values or sense.

The result: expanding the boundaries of the field of cognition – creating new dimensions of sense and axiology.

e. Reflection (R) – meta-cognition and ethical control

Objective: to analyze the process of cognition itself; to distinguish its motivations, limitations, and effects.

Examples:

- The philosophy of science (Popper, Kuhn, Feyerabend) studies how knowledge arises and falls;
- Technology ethics—e.g., questions about responsibility for algorithms' decisions;
- Data quality control and AI audit (AI Ethics Audit);
- Spiritual and philosophical "examinations of conscience" of civilization (e.g., the integral ecology of Pope Francis).

The result: the cognitive system's self-awareness, which allows it to limit its own mistakes and abuses.

Mathematical model of cognition goals

Each act of cognition can be described as a vector in the quantum space of information:

$$|\Psi_p\rangle = \lambda_1 |I_+\rangle + \lambda_2 |I_-\rangle + \lambda_3 |T\rangle + \lambda_4 |X\rangle + \lambda_5 |R\rangle \tag{16}$$

where:

λ_i are the amplitudes of the probability of achieving a given cognitive goal, and their sum meets the condition:

$$\sum_{i=1}^5 |\lambda_i|^2 = 1 \tag{17}$$

Thanks to this, it is possible to formally describe *cognitive equilibrium*—when the amplitude of information and reflection exceeds that of disinformation.

Example: Decision to introduce artificial intelligence into public education

Let's assume that we analyze the cognitive-social process of creating an "AI at school" strategy. Scientists, politicians, teachers, parents, and the media are involved in this process—each with a different cognitive intent. We can assume that the state of cognition of the entire decision-making system can be written as a superposition of the five goals of cognition:

$$|\Psi_p\rangle = \lambda_1 |I_+\rangle + \lambda_2 |I_-\rangle + \lambda_3 |T\rangle + \lambda_4 |X\rangle + \lambda_5 |R\rangle \quad (18)$$

Step 1. Determination of amplitudes (example values)

| The Purpose of Knowledge | Description in the context of AI in education | Amplitude (λ_i) | Probability ($|\lambda_i|^2$) |

I_+ | Informing the public about the opportunities and limitations of AI | 0.6 | 0.36 || | Media disinformation: AI as a "threat to children" or "a miracle solution to everything" | 0.4 | 0.16 || | Transforming Education: New Ways to Teach and Personalize Content | 0.5 | 0.25 || | Transcendence: A Reflection on the Sense of Humanity in the Age of Intelligent Machines | 0.3 | 0.09 || | Reflection: Ethics, Privacy, Right to Error, Data Control | 0.4 | 0.16 | I_{-TXR}

Step 2. Normalization:

$$|\lambda_1|^2 + |\lambda_2|^2 + |\lambda_3|^2 + |\lambda_4|^2 + |\lambda_5|^2 = 0.36 + 0.1 + 0.25 + 0.09 + 0.16 = 1.02 \approx 1 \quad (19)$$

(In practice, minor deviations result from measurement uncertainty—as in quantum physics).

Step 3. Interpretation

- The cognitive system (society, media, and institutions) is in a state of mixed cognition.
- Two goals dominate: information (36%) and transformation (25%), indicating significant development potential.
- However, disinformation (16%) and reflection (16%) balance each other—so the process can easily veer into chaos or dogmatism.
- Transcendence (9%) is low—few people ask *why we need all this*, i.e., about the sense of civilization.

Step 4. Request

The quantum model allows us to describe society as not simply "knowing" what it thinks about AI—it exists in a cognitive superposition until an act of "collapse" occurs (e.g., a government decision, a program reform). The structure of the amplitudes determines whether cognition will lead to wisdom (W) or to a new form of information chaos.

The principle of uncertainty of information

In quantum physics, Heisenberg's uncertainty principle states that it is impossible to determine a particle's position and momentum simultaneously and precisely. The measurement of one of these parameters inevitably affects the other, revealing a fundamental limitation of physical cognition.

$$\Delta x \cdot \Delta p \geq \frac{h}{4\pi}$$

Similarly, in quantum information theory, it can be assumed that information has a pair of complementary attributes: a cognitive goal and a form of transmission. The goal (e.g., informing, disinforming, transforming, etc.) corresponds to the position of the quantum of sense. In contrast, the form (e.g. dialogue, lecture, media, management, preaching) corresponds to its momentum, i.e., the dynamics of influence in the field of communication. Their relationship can be summed up symbolically:

$$\Delta C \cdot \Delta F \geq \kappa \tag{20}$$

Where: means the uncertainty of the cognitive goal, – the uncertainty of the information form, and (κ) is a constant of the quantum field of sense, $\Delta C \Delta F \kappa$ which is the equivalent of the Boltzmann constant in physics. But in *quantum theory, quantum information* has an epistemic-semantic sense, not a physical one. It expresses the minimum threshold of uncertainty of sense—that is, the smallest possible "portion" of uncertainty in the relationship between *the goal of information (C)* and *the form of communication (F)*.

This principle means that you cannot fully control what information means and how it works simultaneously. The more precisely its purpose is defined, the more unpredictable the effect of the form of communication becomes; The more ordered the form, the more open and indefinite the sense remains.

As a result, every act of communication becomes a measurement—a collapse of the function of the sense of information, in which potential senses pass into the reality of interpretation. Ultimately, information is not a fixed state, but a cognitive wave, the amplitude and phase of which depend on the relationship between the sender's intention (goal), the form of communication, and the act of reception (collapse).

The principle of information uncertainty sets the limits of cognition in the digital age: the more we seek to control information, the less truth remains in it. An example of which is the party's or state's proclaimed policy.

Philosophical variant

The same model can be used to analyze historical decisions, e.g.

- Introduction of nuclear energy,
- DNA discoveries,
- or in Poland, to the interpretation of educational reforms after 1989.

In any case, the "vector of cognition" $|\Psi_p\rangle$ can be broken down into ethical, pragmatic, transformational, and disinformation components, and examined to see whether the intellectual balance of a given civilization approaches the state of minimum entropy.

Amplitude of sense and correlation of information quanta

Every act of cognition can be imagined as a wave phenomenon: sense does not appear suddenly but spreads through the mind and culture like an energy wave in space. The amplitude of cognition is encoded in this wave – a measure of the intensity and direction of the sense of information. It is not a statistical number, but *the potential of sense*: semantic energy, which only in the act of interpretation (collapse) becomes measurable.

In quantum physics, amplitude determines the probability of finding a particle in a given state. In quantum information theory, the amplitude of cognition describes the force with which a given. $[\lambda_i]$ cognitive goal – information, disinformation, transformation, transcendence, or reflection – participates in the process of creating sense. When cognition is "measured", that is, interpreted by a human, institution, or artificial intelligence, its potential collapses to its real value:

$$P_i = |\lambda_i|^2 \quad (21)$$

Amplitude, then, is the cause, and probability is the effect; The first is the energy of sense, the second—its trace in the world of facts.

Waves of cognition

Each message, idea, or decision is an information wave with a specific amplitude of cognitive sense. When we speak, write, or analyze data, we produce interference—the overlapping of waves of sense. Two ideas with amplitudes $\lambda_1 \lambda_2$ can strengthen each other (if they are in phase, that is, they agree on the goal) or weaken each other (if they are out of phase).

For example:

- The idea "AI in education increases the availability of knowledge" (I₊) is reinforced with the idea of "AI personalizes teaching" (T),
- But it weakens with the idea of "AI will replace teachers" (I₋).

In this case, cognitive interference occurs³:

$$|\lambda_{1+2}|^2 = |\lambda_1 + \lambda_2|^2 = |\lambda_1|^2 + |\lambda_2|^2 + 2|\lambda_1||\lambda_2|\cos\phi \quad (22)$$

Where:

- ϕ means *the difference in the phases of sense* — that is, the degree of agreement or contradiction of senses.
- $\cos\phi$ - A measure of the interference of sense, from full agreement (1) to complete contradiction (-1).

³ The first person to use the interference formula with the term, was Thomas Young (1801), and in quantum mechanics the full form of this formula was used by Dirac (1927–1930). Which I introduce in relation to wave quantum information. $2|a||b|\cos\phi$

Cognitive interpretation:

- When $\rightarrow \varphi = 0 \cos \varphi = 1$
The phases are consistent \rightarrow constructive interference of senses reinforces each other: knowledge increases, entropy decreases.

Philosophical and cognitive interpretation:

- $\cos \varphi$ expresses cognitive coherence—that is, the degree to which different informational acts interact or collide in the field of consciousness.
- It is the equivalent of *the resonance of sense*: when the phases are in agreement, a new, stronger sense arises (integration of knowledge).
- When the phases are opposite, there is a collapse of sense, loss of sense, or the phenomenon of "cognitive destructive interference" (e.g., propaganda, paradox, contradictory interpretations).

Correlation of information quanta

In classical logic, a message is independent of others. In quantum logic, a message can be entangled with another—its sense depends on the accompanying sense. If two quanta of information are correlated, the state of cognition is not the sum of their states, but a common function: $A B$

$$|\Psi_{AB}\rangle = \lambda_{AB1} |I_+\rangle_A |I_+\rangle_B + \lambda_{AB2} |I_-\rangle_A |I_-\rangle_B + \dots \quad (23)$$

In practice, this means that the sense of one concept depends on the context of the other. For example, the sense of the term "freedom" in the sentence "*freedom to teach*" differs from that of "*freedom from knowledge*". Both information quanta are semantically entangled with each other – you cannot know one without influencing the other.

The correlation of amplitudes is therefore the way in which thought exists in complex social and cultural systems. In the information society, millions of quanta of sense are constantly entangled—through media, algorithms, emotions, and language.

When these correlations are ordered, they form a network of knowledge; when chaotic, information noise. Then the information is no longer a wave with a specific amplitude, but a collection of interfering noises that cancel each other out. Then civilization loses the ability to distinguish between information and disinformation, and thus the direction of cognition.

Measurement and collapsing of sense

The act of interpretation—understanding, deciding, choosing—is the equivalent of quantum measurement. When a human (or AI) performs an act of cognition, the state of superposition of the senses of information collapses: one possibility becomes reality, the others disappear. It is the moment when *possibility becomes a fact and amplitude becomes probability*.

However, cognition does not end at the moment of collapse. Each realized sense of information affects others, changing their amplitudes. In this way, the cognitive system—the individual, society, civilization – is constantly *re-configuring*. It is this non-linear correlation of the quanta of sense that makes the cognitive process never stable: each new piece of information changes the amplitude distribution of all the others.

Cognitive energy and sense coupling

Each quantum of information carries a specific "cognitive energy," which can be approximated through information entropy:⁴

$$S_I = -k \sum_i P_i \ln P_i \quad (24)$$

The more uneven the distribution of probabilities is, the greater the entropy, and therefore the less harmony of cognition. When the amplitudes of information quanta are synchronized, the cognitive system approaches a state of minimal entropy, which can be equated with wisdom. Then the correlation between the quanta of sense becomes positive: information does not compete but cooperates. An example is a scientist or teacher who combines data, ideas, values, and emotions into a coherent model—each element reinforces the other. In contrast, in a system dominated by propaganda or artificial disinformation, amplitudes diverge, leading to destructive interference and cognitive noise. P_i

Towards a quantum ethics of information

A crucial ethical consequence follows from the correlation of the quanta of sense: no act of communication is isolated. Every word, data, decision – even a single "token" – affects the entire cognitive system, modulating the amplitudes of other quanta. Responsibility for a word, science, algorithm, or politics is therefore not a matter of intention, but of the awareness of entanglement.

In this view, wisdom becomes a state of minimal cognitive entropy—an ordering of the senses of information that minimizes conflicts between quanta of information. It is a state in which correlations are harmonious, and amplitudes are in harmony. Then it can be said that *a wise civilization is one in which cognitive interference is constructive rather than destructive*.

Conclusion

The amplitude of cognition is a measure of the potential of the sense of information – the silent energy of thought, which in the act of

⁴ This is due to Claude Shannon's 1948 design.

interpretation turns into fact, decision, knowledge, or wisdom. When the quanta of information are correlated with each other, the world becomes understandable; When the amplitudes of their sense enter destructive interference, chaos ensues. By measuring the amplitudes of cognition, we are not only measuring the power of information—we are measuring *the state of civilizational consciousness*. Because just as a physicist studies the waves of matter, the man of the 21st century must learn to study the waves of sense.

QUANTUM THEORY OF FORMS OF INFORMATION

Category: form of informing

Each act of communication can be assigned to one of six primary forms of information – differing in structure, intention, sender-receiver relationship, and "wave depth" of the message:

CATEGORY	EXAMPLES	DOMINANT FEATURES
A. Dialogic	Conversation, debate, consultation	exchange of amplitudes between peers
B. Monologue	Lecture, sermon, speech	Emission of strong, one-sided amplitude
C. Systemic	Management, Reports, Manuals	Structured, low-noise procedural information
D. Media	Press, TV, Internet, social media	Multi-source interference and chaotic superposition
E. Reflective	essay, meditation, art, poetry	Inner entanglement of senses – introspective information
F. Transcendent	Sermon, Prayer, Moral Manifesto	A message addressed to sense, not only to reason

Quantum characteristics of forms of information

Each form of information can be described as *a quantum state of the field of sense* of information with specific parameters:

- amplitude ($|\lambda|$)—the strength of the message,
- Phase (f)—specific purpose of the message (informing, disinforming, transformation, transcendence, reflection)
- form (r)—a specific form of communication (e.g., dialogue, monologue, system, media, reflection, transcendence, etc.)
- entropy (S)—the degree of chaos or order in a structure,
- Correlation (κ)—the ability to connect with other quanta of information.

Below is an overview of these parameters for each format:

A. Dialogic form—amplitude interference

Explanation: Two or more waves of sense overlap each other. Dialogue is a space of cognitive interference: amplitudes can amplify (when the interlocutors resonate) or abolish (when they exclude each other).

Quantumly, this can be written as follows:

$$|\Psi_{dialog}\rangle = \lambda_A |\phi_A\rangle + \lambda_B |\phi_B\rangle \quad (25)$$

Interference creates a new state of sense that cannot be assigned to either party separately. Dialogue lowers cognitive entropy—if amplitudes are in the realization phase of sense at $\varphi \approx 0$ —and increases semantic correlation κ .

Conclusion: Dialogue is the form closest to quantum entanglement in human communication—it generates a shared sense of information beyond the individual ego.

B. The monologue format—cognitive collapse

Description: The sender sends a strong information impulse – a lecture, a sermon, a speech. The recipient does not interfere but *measures* the message's collapse to its own sense.

Quantumly, this can be written as follows:

$$\hat{M}_{recipient} |\Psi_{speaker}\rangle \rightarrow |\phi_{interpretation}\rangle \quad (26)$$

A high amplitude of the sender ($|\lambda|$ large) means high energy of the message. Still, the lack of feedback increases the risk of losing the realized sense (φ) – the message can "dissipate" in different interpretations.

Conclusion: A monologue is an intense wave, but short-lived—effective in transmitting, weak in consolidating wisdom.

C. Systemic Form—Quantum Order

Description: In management, administration, and science, procedural information dominates, in which sense is controlled by structure (scheme, regulations, algorithm).

Quantumly, this can be written as follows:

$$|\Psi_{sys}\rangle = \prod_i \hat{O}_i |\phi_i\rangle \quad (27)$$

where \hat{O}_i are the operators ordering the information quanta (the components of the system)? Such a message minimizes entropy, but also reduces amplitude—the sense becomes predictable, but not creative.

Conclusion: The systemic form stabilizes information but suppresses creative fluctuations—it is necessary for order, but not sufficient for development.

D. Media format—chaotic superposition

Explanation: In media and social networks, news waves superpose billions of signals, each with a small amplitude, but the combined effect is enormous.

Quantumly, this can be written as follows:

$$|\Psi_{\text{media}}\rangle = \sum_i \lambda_i m^{i\phi_i} |\phi_i\rangle \quad (28)$$

This formula provides a complete, quantum description of the media field of information transmission — that is, how different messages, styles, and communicative intentions affect the *wave of sense-making* in public space. Let us explain step by step what the pointers and elements of this equation mean, especially the power of the form of $m^{i\phi_i}$.

Structure of the equation

It is an extension of the quantum state of media information in the database, where:

- $|\Psi_{\text{media}}\rangle$ — it is the state of the entire media field, the sum of all possible formats and goals of communication,
- λ_i — information amplitude (strength, share, energy of a given channel or form of communication),
- $m^{i\phi_i}$ — a factor describing the modality of transmission (moment, tone, emotion, or context) of the message,
- $|\phi_i\rangle$ — the state of the realized sense of information — thanks to a specific form of communication (e.g., dialogue, monologue, system, etc.).

What does power mean $m^{i\phi_i}$

An expression is a complex form of cognition, fundamental in quantum mechanics and adapted in this theory of the sense of information⁵.

By definition:

$$m^{i\phi_i} = \cos \phi_i + i \sin \phi_i \quad (29)$$

It follows that each component of the message has two coexisting parts:

- real ($\cos \phi_i$) – representing rational, logical sense,
- imaginary ($\sin \phi_i$) – representing the emotional, symbolic, or metaphorical component of the message.

Together, they create the modality of the message, i.e., its *cognitive intonation*.

The importance of modality in the media field

The realized sense of information ϕ_i determines the cognitive moment and the direction of the resonance of the sense because of the use of a specific format of communication:|

⁵ Euler created the mathematics of the complex phase, Schrödinger introduced it to the wave function, and Dirac gave it a fundamental quantum sense.

- Two messages with a similar modality of cognition (e.g., concordant narratives) interfere constructively — they reinforce sense,
- Messages about the opposite modality of cognition (e.g., ideologically contradictory) interfere destructively — they weaken or abolish sense.

That is why the media can create either social resonance (shared sense) or information chaos (decoherence of sense).

Importance of the index i

Index means that the sum runs across all types or channels — examples of i :

- $i=1$: news
- $i=2$: social media,
- $i=3$: science
- $i=4$: religion
- $i=5$: art, etc.

Each has its own amplitude. λ_i and realized sense ϕ_i .

Cognitive interpretation

In quantum information theory, this equation states that the media field is a *superposition of all possible communication modes*, each with its own modalities and amplitudes. The collapse of sense (that is, the moment when a society adopts a single interpretation) occurs when the modalities of sense from different sources are synchronized.

If the phases (goals of cognition) and modalities (forms of communicating media) are random, a decoherence of the sense of information arises — the lack of a typical cognitive structure. Information noise is created, i.e., a state of high cognitive entropy.

Conclusion: The media function as a "quantum field" without measurement — full of potential senses that disappear before they are interpreted. It is the biggest challenge of the digital age.

E. Reflective Form—Internal Interference

Description: Essay, poetry, philosophy—these are the forms in which the sender and receiver are the same consciousness. Information circulates within the cognitive system, producing auto-entanglement.

Quantumly, this can be written as follows:

$$|\Psi_{ref}\rangle = \lambda_{int} |\phi_{mind}\rangle + \lambda_{meta} |\phi_{conscience}\rangle \quad (30)$$

Internal interference lowers the entropy of the sense of information—it leads to cognitive resonance, which can be equated with wisdom.

Conclusion: Reflection is the most stable form of information, with low entropy and high coherence; it is a quantum balance between cognition and being.

F. Transcendent Form—Long Wave Resonance

Description: A sermon, a prayer, an ethical or prophetic manifesto – a message addressed not only to reason, but also to conscience, to a "field of sense" larger than the individual.

Quantumly:

$$|\Psi_{trans}\rangle = \int \lambda(t)m^{i\phi(t)}dt \quad (31)$$

It can be interpreted as a continuous superposition of dynamic states in which amplitude and modality change over time. Let us explain step by step:

Sense of individual ingredients:

- $|\Psi_{trans}\rangle$ —is a "transformed" state (e.g., transformed, transient, or composite), which is the result of summing up all the instantaneous contributions of the state over time.
- $\lambda(t)$ —a function of amplitude (modulus) over time, which determines the strength or intensity of a given instantaneous state.
- $m^{i\phi(t)}$ —a modality factor (forms of information), representing the oscillating nature of information (as in quantum waves or interference signals).
- dt —an element of time, which in the integral means continuous summation after all moments t .

Why does the integral appear?

The integral means that the state $|\Psi_{trans}\rangle$ does not arise in one moment but is the sum of all the momentary contributions—it is as if you were putting together a film of individual frames.

Mathematically, it can be written like this:

$$|\Psi_{trans}\rangle = \text{sum (denoted by integral) of all temporary states } \lambda(t)m^{i\phi(t)} \quad (32)$$

In quantum physics and information theory, this means that the process of transformation (e.g., the flow of information or its energy) extends over time, and its total effect results from the interference of these forms of cognition.

Informative interpretation

In quantum information theory, this can be understood as follows:

- $\lambda(t)$ —the intensity of sense (how much a given moment of cognition brings sense),
- $\phi(t)$ —the realized sense, i.e., its direction of cognition
- Integral – the process of accumulation and interference of senses in time, leading to the creation of the resulting state of information.

In quantum matter, the integral builds a state from the superposition of waves of different frequencies. Here, it builds a state from various formative relationships. If it is a long-length, low-frequency wave, then it is "long-range"

information that can last for generations (e.g., the teachings of Fr. Piotr Skarga, the Gospels, the Decalogue).

Conclusion: The transcendent form is a "standing wave" in the field of sense of civilization—it has the most remarkable ability to perpetuate values.

Interdependence of forms: cognitive interference

In practice, no form of information works in isolation. A conversation (A) can turn into a lecture (B); media (D) reinterpret the sermon (F); reflection (E) becomes a system (C). Their amplitudes interfere, creating a complex cognitive landscape of civilizations.

Mutual reinforcement or extinction depends on the phases of their sense:

COMBINATION

A + E (dialogue + reflection)

B + D (monolog + media)

C + F (System + Transcendence)

A + D (dialog + media)

E + F (Reflection + Transcendence)

INTERFERENCE EFFECT

Cognitive resonance – wisdom is born

Reinforces information chaos

Creates ethical institutional governance

Generates polarization and noise

Leads to a culture of deep sense

Conclusion: form as a quantum state of sense

In quantum information theory, the form of information is not just a carrier of content—it is *an operator of cognition* that changes the state of information, its energy, and direction. Each way of speaking, writing, teaching, or managing acts as a quantum filter that modulates the amplitude of sense and its moral phase.

So we can say:

The quantum phase of information concerns content, but the form organizes the sense of information. And in the quantum age of information, form becomes *somewhat equivalent to content* – both are interdependent quanta of cognition. This assumption confirms the intuition of McLuhan, who in the 60s of the twentieth century announced the famous sentence:

"The medium is the message"

McLuhan noted that it is not the content itself, but the medium—the form of communication—that shapes our way of thinking, our perception of time and space, and even the ethics of communication. Today, we can say that he was right in the quantum sense. Each medium has its own amplitude of sense (having energy), i.e., a characteristic cognitive wave that affects how we perceive reality.

- The book generates long, coherent waves of sense (low entropy, high wisdom).
- Television—short and intense waves (high amplitude, low reflectivity).

- Social media—billions of tiny quanta of sense in superposition (maximum entropy).

As a result, this medium shapes the amplitude of cognition—that is, it decides whether the information becomes knowledge, emotion, illusion, or wisdom. Thus, McLuhan predicted that in the future (i.e., today) form and phase (content) would become inseparable quantum states of communication: one does not exist without the other, and each influences the phase and correlation of the other. In this sense, the modern "quantum theory of information" develops his thought:

- Where McLuhan spoke of a *medium*, we speak of a *wave of information format*;
- Where he saw the *transmitter of culture*, we see the *operator of sense*.

The evolution of the goals of cognition in civilization

AGE	THE DOMINANT GOAL OF COGNITION	CHARACTERISTICS
Antiquity (Socrates, Aristotle)	Reflection (R)	Knowledge for virtue and the common good
Middle Ages	Transcendence (X)	Getting to know God and the sense of existence
Modernity (Bacon, Newton)	Notice (I ₊)	Empirical and applied cognition
Twentieth century (technology)	Transformation (T)	Innovation and control over nature
The 21st century (digital age)	I ₊ -I ₋ hybrid	Coexistence of knowledge and disinformation, the need for reflection (R)

From 2040 to 2050, the goal of cognition is likely to be integrated wisdom (W)—the process that combines all five goals of informing into a dynamic balance.

Towards quantum epistemology

Just as quantum mechanics allows us to predict new states of matter, quantum information theory enables us to predict and organize new states of sense. The information may:

- Interfere (connecting senses),
- Tunnel (jump between contexts),
- To get entangled (to create networks of interdependent ideas),
- To collapse (to realize oneself in interpretation).

In this way, cognition becomes a quantum process—relational, probabilistic, and intentional.

IMPLICATIONS FOR AI

Generative AI works in the spirit of quantum epistemology:

- It operates on tokens as superpositions of senses,
- Uses contextual probability features,
- It generates sense through semantic collapse.

Consequently, humans and AI coexist in the same *field of quantum information*, differing only in motivation: ethical (human) or statistical (AI).

Implication: From Quantum Matter to Quantum Sense

The quantum theory of information shows that information—like matter—has a *dual nature*: wave and particle, potential and actual, chaotic and ordered. The path from data to wisdom reflects the transformation from indeterminacy to stability. Wisdom is a state of minimal cognitive entropy – a balance between knowledge and values.

Integrating quantum matter with quantum information by integrating the cognitive goal with the information format

In classical theories of communication, information is described as content transmitted through a specific channel. In the quantum theory of the sense of information that I develop, information is not just a package of data, but an epistemic state that arises through the interference of two factors: the cognitive goal and the communicative format. Only their combination—mathematically described as a tensor product—creates a complete form of a message, capable of the collapse of sense.

Goal has a phased nature of information

Every communicative act is embedded in a specific purpose of the sender: to inform, persuade, interpret, mobilize, entertain, provoke, or legitimize. The goal is therefore not a side parameter, but a phase function that gives the message a cognitive direction.

A vector describes the goal in phase space:

$$|\kappa_g\rangle,$$

Where the index (g) is the type of goal, and the format factor:

$$m^{i\phi_g}$$

It expresses *cognitive modality*, i.e., the attitude of sense that determines the subsequent interference of information.

The goal models the intention—that is, "what the message is supposed to do" with the mind of the recipient.

Format as a supporting infrastructure of sense

At the same time, there is an independent dimension of the format: monologue, dialogue, debate, column, reportage, panel, academic narrative, emotional narrative, etc. In this theory, format is not a stylistic accessory, but a supporting infrastructure that shapes the course of the collapse of sense. The format is represented as a vector:

$$|\lambda_f\rangle$$

Where index (f) denotes the type of form. On the other hand, the amplitude (λ_{gf}) corresponds to the coupling strength of a given goal with the selected form.

The format determines how the sense is generated and what interpretive course is appropriate for a given message.

Tensor product of goal and format

The whole state of quantum media information arises only when the two dimensions are tensor-connected:

$$|\Psi_{\text{media}}\rangle = \sum_{g,f} \lambda_{gf} m^{i\phi_g} |\kappa_g\rangle \otimes |\chi_f\rangle. \tag{33}$$

Tensor(\otimes) means:

- A multiplication operation that creates a **new space**, larger than the two previous ones.
- Purpose and form are independent dimensions,
- Their combination is not a simple multiplication,
- A new space of communication states is created,
- Each message is a combination (goal, form), not a reducible value.

It is thanks to the tensor structure that it is possible to describe cognitive interference—a phenomenon in which the duplicate content takes on different senses depending on the relationship between goal and form. As illustrated in Table 1.

Table 1. Example of tensor multiplication by formula [33]

Objectives\ Forms	Monologue	Dialogue	Debate	Negotiation / Deliberation
Inform	Information monologue (announcement, announcement)	Information dialogue (interview, Q&A)	Factual debate	Working briefing (reconciliation of facts)
Explain/Teach	Explanatory lecture	Explanatory tutoring/mentoring	Problem seminar	Educational arrangements Workshop
Convince	Persuasive speech	Persuasive conversation	Persuasive debate	Persuasive negotiations
Provoke / Mobilize	Mobilizing speech	Mobilizing/provocative dialogue	Provocative debate	Mobilizing prognosis (call to action)

Interference of Sense as an Effect of the Coupling of Two Spaces

Since the format amplitude has the form:

$$m^{i\phi g}$$

Different goals can interfere constructively or destructively, just as the waves in classical Young's interference or the quantum amplitudes in Dirac's formalism.

As a result:

- The exact words can carry different senses,
- Different forms can strengthen or weaken intent,
- The viewer experiences a collapse of sense dependent on the interference of both dimensions.

It explains why: A persuasive monologue differs from a persuasive dialogue—informing in the form of a debate has a different effect than informing in the form of a reportage—provocation in the form of a dialogue works differently than provocation in the form of a monologue.

Significance for Modern Media

In an age of message overload, the role of purpose and format becomes crucial. It is not the content, but *the modality and infrastructure of the message* that determines what sense will be interpreted.

This theory of quantum information of sense indicates that media do not only transmit information—they transmit cognitive states whose sense is the result of interference between:

- the goal phase,
- the infrastructure of form,
- amplitude weight,

- and the context of the recipient.

Therefore, the analysis of the message must include both dimensions simultaneously.

Update of Einstein's equation: since information is a form of energy

Modern knowledge indicates that matter and information are not separate entities, but two forms of the same energy—as Einstein wrote, "*energy and mass are equivalent*". ($E = mc^2$). In terms of quantum information theory, it can be added that this energy also has an informational dimension, which leads to the equation:

$$E = mc^2 = I k_q T \tag{34}$$

where:

- I means information contained in a system (measured in quantum units of sense),
- k_q — quantum information constant (analogous to Boltzmann's constant),
- T — information time and other (as below), i.e., the period of his dynamic cognitive activity.

What does T mean?

- $T =$ Information Transformation Time (Sense Time). T is the temporal dimension of transforming information into sense, that is, the period during which information *acts, organizes, or changes the state of a system*. It is epistemic time, not physical time. It answers the question: How long does information act to produce energy of sense.
- $T =$ Information Temperature (Processing Intensity).
Where:
 - the degree of stimulation of the information system,
 - the intensity of information exchange,
 - the speed of "boiling sense",
 - Then: $I k_q$ is the informational equivalent of free energy in physics.
- $T =$ Evolutionary dimension (direction of change of information). It is the fourth dimension of quantum information theory—the direction in which information:
 - Grows,
 - Fades,
 - Organizes itself,
 - Collapses into sense.

Why is T needed?

For I (information intensity) to pass into energy, *duration*—processuality—is needed:

$$E_{\text{info}} = I k_q T \quad (35)$$

Without T, information would be "dead", unable to work, untranslatable into the energy of sense. T gives dynamics to information, just as time gives momentum to motion.

Towards the Theory of Everything (ToE)

With the help of cybernetic modeling, we will develop the architecture of the Universe to arrive at a theory of everything by integrating quantum matter with quantum information. The cybernetic, as a graphical model, enriches the existing research of astrophysicists, which is based on mathematical modeling of a fixed composition of four material forces: weak, strong, gravitational, and electromagnetic. On the other hand, graphical modeling highlights the shortcomings of the existing modeling. It detects other forces and the role of the info-communication process.

The model in Fig. 2 highlights the main forces that govern the world and the roles of intelligent control of these forces through info-communication processes.

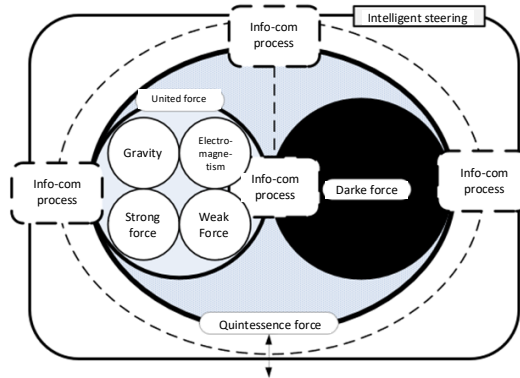


Figure 2. The role of intelligent control of the information and communication process is controlling the operation of the seven forces of the world.

On the other hand, the model in Figure 3 synthesizes these forces and shows that there are 7 of them, i.e., 3 more, which astrophysicists have been operating with so far. It detects the unifying, quintessential, and dark forces. Therefore, their efforts to define a GUT (the Great Theory of Unification) that would integrate these four material forces are in vain.

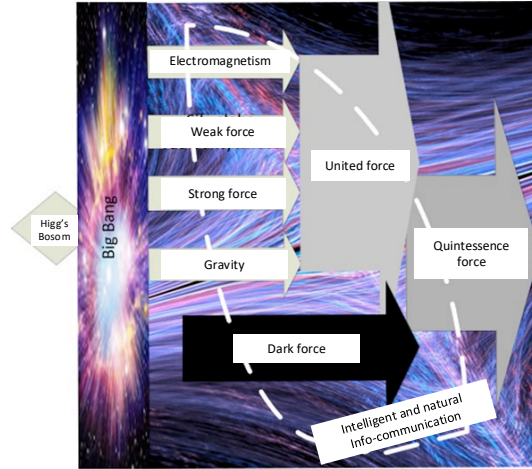


Figure 3. The dynamics of the seven forces of the Universe, which form the basis of GUT (Grand Unified Theory = Grand Theory of Unification), when integrated with the info-communication sense, lead to TOE (Theory of Everything).

From the considerations so far, it follows that *every material system has an information potential*, which manifests as a quantum field of sense. In the context of Figures 2 and 3 of the "Architecture of the Universe," the dynamics of the seven forces of nature can be written as a function of information energy:

$$F_i = -\nabla\Phi_i + \frac{dI_i}{dt}, \tag{36}$$

Where:

- F_i - the i^{th} force of the Universe (e.g., gravity, electromagnetism, weak, strong, dark, quintessence, informational force),
- Φ_i - the potential of the field,
- $\frac{dI_i}{dt}$ - a stream of information affecting the change in the state of energy.

In dynamic equilibrium, the sum of all forces balanced by information flows is (in vector notation⁶):

$$\sum_i \vec{F}_i = \frac{d\vec{I}_{\text{tot}}}{dt} = 0 \tag{37}$$

⁶ Writing an equation in vector notation means that all the quantities in this equation have a direction and a turn—they are not ordinary numbers (scalars), but vectors in space. So they must be summed and differentiated as vectors – and equality to zero means full directional equilibrium in the state space.

Which describes the state of information-material equilibrium – the equivalent of the state of minimal entropy in the theory of cognition. When $\frac{dI_{tot}}{dt} > 0$, there is an expansion of the Universe (expansion of the information field), and when $\frac{dI_{tot}}{dt} < 0$ – it's the contraction of the Universe or collapse of sense.

In terms of information, it can be read as follows:

- \vec{F}_i – are local "forces of sense" (factors that change the direction or intensity of sense in a system, under a vector notation),
- \vec{I}_{tot} – total vector of information flow throughout the system,
- The equation states that when the sum of all sensory-creative interactions is balanced, the information system reaches a state of stable interpretation—no new senses arise or disappear.

It means that the total impulse (or momentum) of the information is constant, and the tot index (from *total*) indicates the sum of all components of the system—the total state, not the local state.

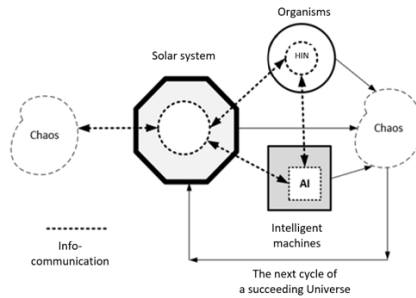


Figure 4. Universe Architecture Series (HIN-Higher Intelligence, AI-Artificial Intelligence)

The cybernetic model in Figure 4 of the Universe Architecture Cycle shows the *role of info-communication control* as a process that synchronizes the flow of energy and information between quantum matter systems. Formally, this can be summed up by the equation of information control:

$$\frac{dI}{dt} = \alpha \nabla \cdot (c \Psi) \quad (38)$$

Where: Ψ is the information wave (a function of the state of sense), and α – the coupling factor of intelligent control (equivalent to the system's ability to self-regulate). Growth of α corresponds to the increasing coherence of the Universe—a state in which information and matter are synchronized, and physical and cognitive phenomena become one function of evolution. c = the *rate, velocity, or intensity* with which the information field Ψ spreads through space.

Understood in this way, the integrated theory of quantum matter and quantum information creates a generalized version of Einstein's theory – covering not only energy and space-time, but also information as the third dimension of existence. It can be written symbolically:

$$E^2 = (mc^2)^2 + (Ik_qT)^2 \tag{39}$$

It expresses the coexistence of material and informational energy in a single, non-linear structure of the Universe. In the state of information equilibrium, both terms are harmonized ($E_m = E_l$), while in the state of expansion ($E_l > E_m$). There is an expansion of space, according to Einstein's and Hubble's cosmological observations, but extended by the cognitive dimension. This formula extends Einstein's relativistic energy formula, and the squares on both sides have profound physical and structural significance. Let's explain step by step.

Classic pattern form

Einstein's equation for the total energy of a (relativistic) system is:

$$E^2 = (mc^2)^2 + (pc)^2 \tag{40}$$

Where:

- E —total energy,
- m —rest mass,
- p —momentum,
- c —the speed of light.

This equation combines the resting energy (mc^2) with the energy of motion (pc) into a single geometric whole –like Pythagoras' theorem:

$$E^2 = E_{\text{resting}}^2 + E_{\text{motion}}^2 \tag{41}$$

That is why there is a square form—to combine quantities of different types (energy and momentum) in a geometrically consistent way.

In the quantum version of the information:

$$E^2 = (mc^2)^2 + (Ik_qT)^2 \tag{42}$$

The second component is the quantum-information analogue of momentum: (Ik_qT)

- I — information intensity (information pulse),
- k_q — the quantum constant of information (analogous to k_B),
- T — time, temperature, or period of transformation of sense.

Therefore:

$$E^2 = (mc^2)^2 + (Ik_qT)^2 \tag{43}$$

It plays the role of a dynamic component of information energy, just as pc it plays the role of kinetic energy in relativistic physics.

Why are both sides squared?

(a) For geometrical reasons.

The equation has a Euclidean (Pythagorean) structure in four-dimensional space [time–total energy–mass energy (mass being), information energy (sense being)]. Squares allow you to "add" quantities that differ in direction or dimension—the total energy is then the length of the vector in the energy space:

$$E^2 = E_{\text{resting}}^2 + E_{\text{inform}}^2 \quad (44)$$

Without squares, it would not be possible to add these quantities.

(b) For physical reasons

Squares eliminate negative signs and ensure that:

- energy is always a positive quantity (regardless of the direction of information flow),
- time symmetry is maintained (both for $+T$ and $-T$)
- The duality between mass energy and information energy can be described without a unit collision.

(c) For information purposes

In quantum information theory, squares mean:

- real (measurable) values of the energy of sense, because the moduli of the squares of amplitudes are always observable,
- Thus E^2 it corresponds to the "density of sense," just as $|\Psi|^2$ it corresponds to the probability density.

General interpretation

$$E^2 = (mc^2)^2 + (Ik_q T)^2 \quad (45)$$

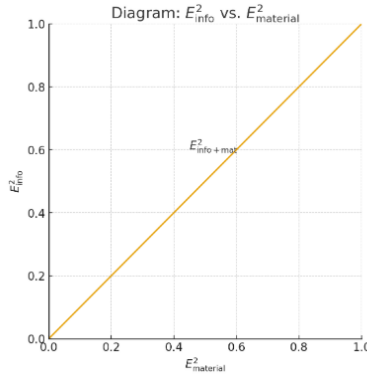


Figure 5. Summation chart of material and informational energy (E_{tot})

can therefore be read as:

The total energy of being (or sense) is the geometric sum of two components:

- energy encoded in matter (mass),
- and the energy encoded in the flow of information (the impulse of sense, time, the temperature of cognition).

Metaphorically:

Just as in physics, matter and motion are two sides of the same energy, so in quantum information theory, matter and information are two forms of the same quantum reality of sense, and their combination requires square geometry.

Conclusion

1. The Universe is an information-material system in a state of dynamic equilibrium

From the equation (in vector notation):

$$\sum_i \vec{F}_i = \frac{d\vec{I}_{tot}}{dt} = 0 \tag{46}$$

It follows that matter and information are two aspects of the same energy of existence. The balance of physical forces (gravitational, electromagnetic, etc.) is only possible thanks to the flow of information that synchronizes it. When this flow is positive, the Universe expands—both cosmologically and cognitively; when negative, there is a collapse of energy and sense.

Conclusion: The Universe maintains stability not only through the laws of physics but also through the flow of information – it is a self-regulating information-energy system.

2. Information acts as an equivalent agent of energy and matter

Extended Energy Identity Equation:

$$E^2 = (mc^2)^2 + (Ik_qT)^2 \tag{47}$$

It shows that information has an energetic dimension and can affect physical and cognitive processes. Increasing informational order (coherence) raises the level of organization of matter, from atoms to biological and social systems.

Conclusion: Information is not just a description of the world, but its active component — it co-creates matter, energy, and consciousness.

3. The evolution of the cosmos and man has a standard information code
The equation describes info-communication control:

$$\frac{dl}{dt} = \alpha \nabla \cdot (c\Psi) \quad (48)$$

Shows that the Universe develops through an increase in the coefficient of coherence α — that is, the ability to intelligently self-organize. In the biological and civilizational dimension, this corresponds to the development of consciousness, wisdom, and the ability to cooperate.

Conclusion: The growth of wisdom in cognitive systems (human (WIN), AI, civilization) is a continuation of the cosmic process of ordering information — the evolution of spirit is a natural extension of the evolution of matter.

How the energy of the material quantum is added to the energy of the information quantum

Example with NVIDIA shares.

In the Theory of Everything, the energy of every phenomenon has two components:

1. Material energy – related to real, measurable parameters of the system (here: capitalization, revenues, technology, products).
2. Information energy
– related to the influx of information that changes the interpretation of the value of the system (here: government decisions, licenses, regulations, market narratives).

Together they form:

$$E_{\text{tot}} = E_{\text{mat}} + E_{\text{info}} \quad (49)$$

In the financial market, NVIDIA's capitalization is an example of material energy—it results from real products, the computing power of its GPUs (Graphics Processing Units, i.e., specialized computing chips designed to perform parallel AI calculations very quickly), and actual market-driven demand. It is the energy of "mass" that exists independently of interpretation.

However, at some point, information emerges: the government is limiting export licenses and imposing high tariffs on imports of digital components. This information acts as a *quantum impulse of sense*: it changes investors' perceptions and shifts the market into a new state. Although it is non-material

information, it has energy — it transforms expectations, business models, and therefore the valuation of shares.

As a result:

$$E_{\text{tot}} = E_{\text{mat}} + E_{\text{info, (regulations, market)}} \quad (50)$$

The value of stocks is rising not because there are physically more chips, but because there is more information that changes the interpretation of future earnings.

It is a classic case where:

- material quantum (real value of products)
- information quantum (government decisions, narrative, flow of sense)

They add up, creating a spike in the system's total energy—that is, an increase in the stock's valuation.

Because financial markets are a pure implementation of the principle of the ToE:

Material energy + Information Energy → Valuation Energy.

The market responds not just to a company's output, but especially to what is known and believed about it.

Gold Example

- The material energy of gold is stable. The information energy of gold is unlimited. And their sum creates its real value.
- So gold is valuable not because it is heavy, but because it carries a gigantic energy that organizes human decisions.

The example of TESLA

- Tesla is a material company only in its factories.
- On the stock exchange, it is an information company.
- And its real energy is the sum of both.
- Example of war in Ukraine
- Ukraine won with information energy (the role of President Zelenskyi), which it could not win with material energy.
- And the West added matter to it thanks to the power of this information.

An example of China's role in the 21st century

- China, as the World Factory
- with the power of narrative and standards and the wisdom of its sages ("win a war without a battle"),
- to co-create the global rules of the 21st century.

Most studies on human longevity point to a particularly important, if not the most important, factor [e.g., according to Japanese studies (ikigai harmony) and the Harvard Study of Adult Development (Vaillant 2012)]: having a sense

of life and of the next day. This kind of sense strengthens the energy of life towards the activity of actions and relationships, which prolongs life.

How quantum information can develop the queen of sciences, which is philosophy, i.e., the love of wisdom, the highest possible sense of people's lives.

First, we will give its quantum definition:

Philosophy is the process of harmonizing the amplitudes of sense in the search for wisdom – the love of order in the information field of the world.

With *Homo sapiens*, we may become *Homo sapientiae harmonicus* = a *Man of wisdom and balance*, who can transform information into order, knowledge into compassion, and technology into a tool of development, but not destruction. Therefore, philosophy can develop in the following ways:

i. Metaphysics and ontology

- Relational beings (ontology of states, not substances): we describe what exists as a *network of information states* and their correlations (entanglements), not as "self-contained" things.
- Informational monism/processualism: events are "collapses of sense".
- Potential layer: the superposition of states gives rise to an ontological space of possibilities, not a "lack of knowledge". The philosophy of possibility (Aristotle) regains its formal apparatus.
- *Epistemology*
- Cognition as a measurement: The cognitive act acts as an operator: $\widehat{M} - \square$ context and *intention* co-create the outcome. The classic ideal of "knowledge from nowhere" gives way to precise contextualism (the so-called "*connecting the dots*").
- Probability of sense: the thesis introduces $P_i = |\lambda_i|^2$. The rationality metric allows us to evaluate and contrast the "energy" behind different cognitive goals, such as informing, misleading, transforming, transcending, and reflecting.
- *Philosophy of mind and consciousness*
- Correlation instead of location: We model mental content as patterns of information correlation (not "representations in a box").
- Naturalization of intention: The amplitudes and phases of "cognition goals" can give a formal description of intentionality (why the thought is "directed at something").
- *Logic and methodology*
- Multivalued/quantum logics: formalize *the superposition of truth* before interpretation, while truth in ToE appears as a limit (after the collapse).

- Probabilistic-operational method: humanities studies can use vectors of states of sense and normalization to compare discourses, narratives, and doctrines.
- *Philosophy of science*
- Complementarity of paradigms: classical and quantum models are *complementary* descriptions of the same informational reality.
- Forecasting "states of information": just as we predict states of matter, we can design states of information (ecosystems of knowledge, public policies, education, and so forth).

New theses that philosophy can adopt

1. The thesis of information correspondence: truth = high correspondence of distribution of sense between models, data, and practice (coherence + pragmatics).
2. The wisdom thesis of minimal entropy: wisdom is the ability to organize the field of sense in such a way as to reduce chaos without suppressing pluralism.
3. The thesis of entangled responsibility: each word/algorithm changes the amplitudes of the others—the ethics of communication becomes the ethics of *interference*.

CONCLUSION

- If the twentieth century quantized matter, then the twenty-first century faces the task of quantizing sense, understood as a structural property of information systems. It means a shift in philosophy towards *information physics*, in which the waves of sense are subject to the same laws of interference, superposition, and collapse as the waves of matter.
- Understanding cosmic evolution requires analyzing the "black box" of the Universe, i.e., those processes in which local gradients of information and fluctuations in entropy generate increasingly complex configurations. Complexity here is not an accidental result of chaos, but the result of *the action of informational organizing fields* that direct the system towards states of higher coherence.
- The first living cell was created through a sequence of information transformations in an environment with a high density of chemical data. Analysis of these processes indicates that life originated where channels with minimal noise appeared, capable of maintaining and reproducing patterns – in other words: where there was *a stable flow of information*.
- The processes of Darwinian natural selection can be reinterpreted in terms of information theory as mechanisms for code optimization and entropy reduction. The apparent randomness of mutations is compensated for

by systemic pressure to increase the body's informational efficiency, leading to structures of high complexity and coherence stability.

- The question arises whether there is a form of external or emergent control of the flow of information, analogous to *the operators of evolution* in quantum mechanics [e.g., the Higgs boson supports the Standard Model because it proves that particles gain mass (almost at 0 time⁷) from the Higgs field, precisely as the theory predicted]. *In the Theory of Everything, the Higgs mechanism can serve as the bridge linking quantum fields of gravitation and information by treating mass-generation as an emergent, energy-informational process embedded in the Universe's fundamental fabric.* It may even imply the existence of higher-order rules governing the direction of sense propagation in biological, technological, and social channels.

- The transformation of life cannot be fully explained in terms of classical matter, as it requires a model that accounts for both energy and information flows. Information acts as *a control field, reducing chaos and maintaining the system's structural order.*

- The use of the term "control" in the context of information physics suggests the existence of signals with a specific channel capacity, bandwidth, and noise immunity that determine the evolution of a system. These signals act as design operators, directing the flow of content towards states with greater stability of sense.

Therefore, a new paradigm of the architecture of the Universe is emerging, in which matter, energy, and information should be treated as conjugate field variables and regarded as emergent properties of systems with high informational coherence. This paradigm may become the foundation of a new science integrating physics, quantum computing, and epistemology in the coming centuries.

REFERENCES

- I. Asimov, *Science Fiction Works*, Various editions.
 D. Bohm, *Quantum Theory*, Routledge, London 1952.
 D. Bohm, *Wholeness and the Implicate Order*, Routledge, London 1980.
 D. Bohm, *Thought as a System*, Routledge, London 1992.
 D. Bohm, B. Hiley, *The Undivided Universe: An Ontological Interpretation of Quantum Theory*, Routledge, London 1975.
 N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford 2014.
 A. C. Clarke, *Profiles of the Future*, Harper & Row, New York 1962.
 R. Dawkins, *The Blind Watchmaker*, W. W. Norton, London 1986.
 R. Dawkins, *The God Delusion*, Bantam Press, London 2006.
 L. de Broglie, *Introduction à la Mécanique Ondulatoire*, Gauthier-Villars, Paris 1928.
 H. von Ditfurth, *The Origins of Life: Evolution as Creation*, Harper & Row, New York 1982.
 A. Einstein, *The World as I See It*, Philosophical Library, New York 1934.

⁷ Some call it the God's particle since is almost at -T, but "almost".

- L. Floridi, *Philosophy of Information*, Oxford University Press, Oxford 2011.
- V. Flusser, *Towards a Philosophy of Photography*, Reaction Books, London 1983.
- B.-C. Han, *In the Swarm: Digital Prospects*, MIT Press, Cambridge (MA) 2017.
- S. Hawking, *A Brief History of Time*, Bantam Books, London 1988.
- P. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Physical Review Letters, 13(16), 1964.
- E. Hubble, *A Relation Between Distance and Radial Velocity Among Extra-Galactic Nebulae*, Proceedings of the National Academy of Sciences, 1929.
- L. M. Krauss, *A Universe from Nothing: Why There Is Something Rather Than Nothing*, Free Press, New York 2012.
- H. Krawczyk, A. Targowski, *Wisdom in the Context of Globalization and Civilization*, Cambridge Scholars Publishing, Newcastle 2019.
- R. Kurzweil, *The Singularity Is Near: When Humans Transcend Biology*, Viking, New York 2005.
- J. Lennox, *God's Undertaker: Has Science Buried God?*, Lion Books, Oxford 2009.
- A. Loeb, *Extraterrestrial: The First Sign of Intelligent Life Beyond Earth*, Houghton Mifflin Harcourt, Boston 2021.
- M. McLuhan, *Understanding Media: The Extensions of Man*, McGraw-Hill, New York 1964.
- W. J. Ong, *Orality and Literacy: The Technologizing of the Word*, Methuen, London 1982.
- R. Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe*, Jonathan Cape, London 2004.
- N. Postman, *Amusing Ourselves to Death*, Viking, New York 1985.
- C. Sagan, *The Demon-Haunted World: Science as a Candle in the Dark*, Random House, New York 1997.
- E. Schrödinger, *What is Life? The Physical Aspect of the Living Cell*, Cambridge University Press, Cambridge 1944.
- P. Sloterdijk, *Sphären I–III*, Suhrkamp, Frankfurt am Main 1998–2004.
- C. E. Shannon, *A Mathematical Theory of Communication*, Bell System Technical Journal, 27, 1948, s. 379–423, 623–656.
- C. E. Shannon, W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana 1949.
- B. Spinoza, *Ethica, ordine geometrico demonstrata*, Amsterdam 1677.
- B. Stiegler, *Technics and Time I–III*, Stanford University Press, Stanford 1994–2001.
- A. Targowski, *The Architecture and Planning of Enterprise-wide Information Management Systems*, Idea Group Publishing, Hershey 1990.
- A. Targowski, *Information Technology and Societal Development*, Information Science Reference, Hershey 2009.
- A. Targowski, *The Harnessing the Power of Wisdom*, NOVA Science Publishers, New York 2015.
- A. Targowski, *The Strategies of Informing technology in the 21st Century*, IGI Global, Hershey 2019.
- A. Targowski, *Strategic Informatics*, Polish Academy of Sciences, Kraków 2023.
- A. Targowski, H. Krawczyk, *Wisdom in the Context of Globalization and Civilization in the Age of Universal Digitization*, CapitalBook, Warsaw 2025.
- G. E. Vaillant, *Triumphs of Experience: The Men of the Harvard Grant Study*, Harvard University Press, Cambridge (MA) 2012.
- H. Von Dittfurth, *The Origins of Life: Evolution as Creation*, Harper & Row, New York 1982.
- C. F. Von Weizsäcker, *The Unity of Nature*, Farrar, Straus & Giroux, New York 1974.
- S. Weinberg, *Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature*, Vintage Books, New York 1993.
- A. Wolszczan, *A Planetary System Around the Millisecond Pulsar PSR1257+12*, Nature, 355, 1992.
- S. Zuboff, *The Age of Surveillance Capitalism*, PublicAffairs, New York 2019.
- W. H. Zurek, *Decoherence, Einselection, and the Quantum Origins of the Classical*, Reviews of Modern Physics, 75(3), 2003, s. 715–775.

**KWANTOWA TEORIA INFORMACJI: W STRONĘ PARADYGMATU
POZNANIA, TEORII WSZYSTKIEGO I FILOZOFII**

STRESZCZENIE

Jest to autorska koncepcja *kwantowej teorii informacji* jako uniwersalnego paradygmatu poznania oraz *teorii wszystkiego*, integrującej fizykę, epistemologię i ontologię sensu. Autor zakłada, że informacja nie jest jedynie opisem rzeczywistości, lecz jej *pierwotną materią* – energią poznawczą, która współtworzy materię, świadomość oraz ewolucję cywilizacji. Model opiera się na *pięciostopniowej hierarchii transformacji poznawczej*: dane → informacja → koncepcja → wiedza → mądrość, odpowiadającej stopniowemu porządkowaniu *entropii sensu*. Proces poznania ma *naturę kwantową*: każdy akt myślenia istnieje w stanie *superpozycji sensów*, a dopiero w momencie interpretacji następuje *kolaps sensu*, czyli przejście od potencjalności do poznawczej aktualności. Autor rozwija analogię pomiędzy *falą poznania* a *falą materii*, wprowadzając pojęcia *amplitudy sensu* oraz entropii poznawczej jako formalne miary jakości poznania. Rezultatem jest teoria, w której *człowiek, sztuczna inteligencja oraz Wszechświat* uczestniczą w jednym polu kwantowej informacji, a *równowaga poznawcza* odpowiada stanowi minimalnej entropii – mądrości. Artykuł integruje klasyczne koncepcje *Bohma, Einsteina, McLuhana i Floridiego*, ukazując informację jako czynnik sprawczy równoważny energii i materii. Zaproponowana koncepcja *Homo sapiens harmonicus* przedstawia człowieka przyszłości jako istotę zdolną do harmonizowania amplitud sensu, łączenia wiedzy z wartościami oraz przekształcania technologii w narzędzie rozwoju duchowego. Pracę zamyka postulat *epistemologii kwantowej*, w której filozofia staje się nauką organizowania informacyjnego pola świata – sztuką przekształcania chaosu danych w *mądrość istnienia*.

Słowa kluczowe: kwantowa teoria informacji; epistemologia kwantowa; kolaps sensu; entropia poznawcza; mądrość; filozofia informacji; amplituda poznania; teoria wszystkiego; *Homo sapiens harmonicus*; McLuhan; integracja materii i informacji.

O AUTORZE — Professor Emeritus of Computer Information Systems, at Western Michigan University. Distinguished as a Golden Engineer of the Warsaw University of Technology and as a Golden Engineer of the 30th Anniversary of the IIIrd Polish Republic [organized by the Polish Federation of Engineering Associations (NOT)]

Email: andrew.targowski@wmich.edu

Stanisław Buda

WSZECHŚWIAT JAKO SYSTEM IZOLOWANY

<https://doi.org/10.37240/FiN.2025.13.03>

STRESZCZENIE

Niniejszy artykuł proponuje nowatorski model fizyczno-ontologiczny Wszechświata jako absolutnie izolowanego systemu charakteryzującego się niezerową egzergią, relacyjnym porządkiem oraz ontologiczną prymarnością informacji. Podważając klasyczną termodynamiczną prognozę śmierci cieplnej, argumentuje, że entropia jest zależna od skali opisu, a globalna egzergia jest stale podtrzymywana dzięki dynamicznemu sprzężeniu energii i informacji. Lokalne podsystemy ($P \subseteq W$) wyłaniają się jako samoregulujące centra porządku, koncentrujące egzergię i redukujące entropię informacyjną. Model integruje koncepcje z konformalnej kosmologii cyklicznej (Penrose), zasad holograficznych oraz przyczynowości zstępującej, sugerując holistyczną, idealistyczną kosmologię, w której całość poprzedza części (hologeneza). Implikacje obserwacyjne obejmują reinterpretację anizotropii CMB jako śladów dziedziczonej informacji z poprzednich eonów oraz testy za pomocą Euclid i JWST. Wszechświat nie jest systemem zmierzającym ku równowadze, lecz samopodtrzymującym się procesem relacyjnym — istnieniem jako wieczny ruch, sens i samoregulacja.

Słowa kluczowe: Wszechświat, ontologia, egzergia, informacja, entropia, hologeneza.

1. WSTĘP

Współczesna fizyka i kosmologia oferują spójny opis zjawisk zachodzących we Wszechświecie, ale napotykają trudności, gdy próbują objąć całość bytu jednym modelem. Zagadnienia te dotyczą nie tylko granic empirycznych teorii, lecz także ich podstawowych założeń ontologicznych.

Najistotniejsze problemy tradycyjnych ujęć można ująć w kilku punktach:

1. Problem entropii i tzw. „śmierci cieplnej” Wszechświata. Zgodnie z klasycznym ujęciem drugiej zasady termodynamiki (Clausius, 1850 – wzrost entropii w układach izolowanych) każda izolowana całość dąży do stanu maksymalnej entropii – równowagi, w której zanika możliwość wykonania pracy. Przeniesione na skalę kosmiczną założenie to prowadzi do koncepcji „śmierci cieplnej Wszechświata”, a więc do obrazu rzeczywistości, w której wszelka aktywność ma zaniknąć. Tymczasem ujęcie to zakłada, że entropia jest wielkością absolutną, a nie relatywną wobec poziomu organizacji i sposobu opisu. Pomija też fakt, że porządek może mieć charakter lokalny, a egzergia (zdolność do wykonania pracy w danym otoczeniu) może być rozproszona, lecz niezerowa.
2. Niejasność pojęcia „izolacji” w odniesieniu do całości bytu. W fizyce izolacja oznacza brak wymiany energii lub materii z otoczeniem. W przypadku Wszechświata pojęcie to traci sens operacyjny – nie istnieje bowiem „zewnątrz”. Ujęcia klasyczne nie rozróżniają między empiryczną izolacją a ontologiczną jednością, która jest właściwa całości bytu.
3. Brak adekwatnego modelu relacji między energią, informacją i strukturą. W większości współczesnych teorii energia i materia uznawane są za pierwotne, a informacja – za ich opis. Jednak coraz wyraźniej ujawnia się, że informacja nie tylko opisuje porządek, lecz go wytwarza: ogranicza przestrzeń stanów, ukierunkowuje przepływy energii, a przez to realnie zawęża możliwe konfiguracje systemu. W tym sensie informacja pełni funkcję ontologiczną, nie tylko semantyczną.
4. Trudność wyjaśnienia spontanicznego powstawania porządku. W ramach tradycyjnych modeli termodynamicznych trudno opisać pojawianie się złożonych struktur – biologicznych, poznawczych, kulturowych – bez odwoływania się do przypadkowych fluktuacji lub bez odwoływania się do przypadkowych wahań lub specjalnego dostrojenia warunków na początku. Brakuje pojęcia, które pozwalałoby myśleć o porządku jako wewnętrznej własności systemu, wynikającej z samej struktury jego relacji.

Proponowane założenia

W niniejszym tekście przedstawiam próbę rozwiązania powyższych trudności poprzez model, w którym Wszechświat zostaje ujęty jako system absolutnie izolowany, a jego wewnętrzna dynamika zostaje opisana przy pomocy pojęć energii, informacji i egzergii.

1. Wszechświat jako absolutnie izolowany system. Wszechświat jest całością obejmującą wszystkie możliwe stany, relacje i prawa – sys-

- temem, który nie posiada żadnego zewnątrz. „Izolacja” oznacza tu nie tyle brak kontaktu z otoczeniem, ile brak czegokolwiek poza całością: wszelka relacja zachodzi wewnątrz Wszechświata, a jego jedność polega właśnie na immanentnej pełni tych relacji.
2. Niezerowa egzergia jako warunek ontologicznej aktywności. Wszechświat nie jest układem zmierzającym ku absolutnej równowadze, lecz strukturą, w której globalna egzergia jest zawsze niezerowa. Oznacza to, że zawsze istnieje pewien potencjał zdolny do pracy – do generowania porządku i przemiany. Niezerowa egzergia jest nie tyle parametrem fizycznym, ile zasadą ontologiczną: warunkiem tego, że rzeczywistość w ogóle „działa”.
 3. Informacja jako czynnik zawężający przestrzeń stanów. Informacja nie jest biernym zapisem, lecz realnym czynnikiem ograniczającym. Jej funkcją jest redukcja możliwości – zawężanie przestrzeni stanów do takich, które mogą zostać zrealizowane w danym kontekście. Tym samym informacja stanowi negatywną zasadę porządku: przez ograniczenie umożliwia działanie. W ujęciu fizycznym odpowiada to sterowaniu przepływem energii; w ujęciu ontologicznym – konstytuowaniu sensu.
 4. Relacyjna ontologia porządku. Elementy Wszechświata nie są substancjami, lecz węzłami w sieci relacji. Porządek, stabilność i trwałość struktur wynikają z dynamicznego równoważenia przepływów energii i informacji. Każda forma bytu jest lokalną równowagą tych dwóch czynników – jest utrwalonym wzorem ograniczenia.
 5. Systemowa wieloskalowość i emergencja. Ponieważ informacja i energia pozostają sprzężone, Wszechświat może wytwarzać coraz bardziej złożone struktury poprzez lokalne zwiększanie porządku kosztem entropii otoczenia. Lokalne podsystemy $P \subseteq W$ – od galaktyk po systemy biologiczne – mogą być rozumiane jako ośrodki samosterującej się egzergii: organizmy, które utrzymują swój porządek dzięki przepływowi energii i informacji.
 6. Ciągła aktywność Wszechświata. Model ten eliminuje pojęcie ostatecznego „wychłodzenia” czy zaniku aktywności. Tak długo, jak istnieją relacje i przepływy informacji, egzergia nie znika, lecz przekształca się: porządek może zmieniać formę, ale nie przestaje istnieć. Śmierć cieplna byłaby więc raczej granicznym przypadkiem idealnym, niemożliwym do osiągnięcia w rzeczywistości dynamicznej i relacyjnej.

W dalszych częściach tekstu zostanie przedstawiony formalny model Wszechświata jako systemu izolowanego, wraz z opisem jego wewnętrznych relacji i podsystemów. Celem jest nie tylko zarys nowego paradygmatu fizykalno-ontologicznego, ale także próba odpowiedzi na pytanie, dlaczego rze-

czywistość trwa, działa i wytwarza sens — mimo że z perspektywy czysto termodynamicznej powinna już dawno „zamilknąć”.

2. DEFINICJA WSZECHŚWIATA JAKO SYSTEMU IZOLOWANEGO

Wszechświat traktujemy jako system absolutnie izolowany. Znaczy to, że obejmuje w sobie wszystko, co istnieje; że istnieje on bez przyczyny zewnętrznej; że nie ma sensu mówić o jego początku ani końcu. Jeśli jest systemem, to zawsze musi mieć swoje elementy (dyskretne lub ciągłe), które mogą wchodzić ze sobą w takie czy inne możliwe interakcje, tzn. muszą na siebie wzajemnie oddziaływać. Oddziaływania te sprawiają, że między tymi elementami utrzymują się pewne więzi (w dopuszczalnych granicach). Wewnątrz systemu muszą zachodzić różne dopuszczalne zmiany, modyfikacje, konfiguracje. W systemie może się zmieniać wnętrze, może się również zmieniać on sam (jego własności jako całości). Nie zmienia się natomiast jego natura (tj. strukturalne reguły opisu: przestrzeń stanów i prawa dynamiki). O tym, że wszechświat jest jednością (całością) nie decyduje jego unikalność (jedyność), ale jego holistycznie pojmowana natura.

W języku bardziej sformalizowanym: Niech wszechświat (W) będzie traktowany jako system absolutnie izolowany, tzn. jako całość obejmująca wszystkie dopuszczalne konfiguracje i nieposiadająca zewnętrznego odniesienia ani zewnętrznych reguł działania. System W definiujemy przez uporządkowaną krotkę: $W = (X, \Phi, E, I, \mu, O)$, gdzie:

- X to zbiór możliwych konfiguracji (przestrzeń stanów) — komplet opisów, które system może przyjąć;
- $\Phi: X \rightarrow X$ to reguła przejść (ewolucji) określająca, jakie transformacje konfiguracji są dopuszczalne;
- E to rodzina „elementów” lub podstruktur (dyskretna lub ciągła), przez które można dekomponować konfiguracje systemu;
- $I \subseteq E \times E \times T$ to struktura interakcji — relacje i typy powiązań między elementami oraz warunki ich aktywacji;
- μ to sposób ważenia możliwych stanów systemu — reguła mówiąca, które konfiguracje uważamy za bardziej typowe albo ważniejsze w naszym opisie;
- O to zbiór funkcji obserwowalnych / właściwości agregowanych (reguły ekstrakcji cech), wedle których można formułować opisy makro i oceniać zmiany.

W tym ujęciu „izolacja” W znaczy — formalnie — brak rozszerzalnej krotki $W' \supset W$ z dodatkowymi, zewnętrznymi regułami lub konfiguracjami; „jedność” W wynika nie z bycia jedynym bytem ontologicznym, lecz z sieci relacji I i z tego, że wszystkie dopuszczalne konfiguracje należą do jednej przestrzeni X . „Natura” systemu rozumiana jest tu jako zestaw niezmienni-

ków strukturalnych (X, Φ, I) – zasad, które nie ulegają zmianie w sensie definicyjnym, nawet jeśli jego elementy, konfiguracje lub właściwości agregowane ulegają ewolucji. Tak sformułowany model jest abstrakcyjną ramą opisu: nie narzuca ontologicznych hipotez o tym, czym są „elementy”, ani nie determinuje metody wyboru μ czy O – te wybory zależą od poziomu opisu i od celu analizy.

Zauważmy, że traktowany w ten sposób Wszecławiat pełni rolę systemu przede wszystkim dlatego, że jego części pozostają ze sobą w relacjach i procesach. Wewnętrzna równowaga tak rozumianej całości ma charakter dynamiczny: relacje między elementami nieustannie się rekonfigurują w dopuszczalnych granicach, a nawet gdy cechy makroskopowe wydają się stałe, to najczęściej są one wynikiem ciągłych mikrozmian skompensowanych na wyższym poziomie. Teoretyczna możliwość absolutnego braku wszelkich zmian (absolutny zastój na wszystkich poziomach) opisywałaby raczej statyczną strukturę niż system czynny – w sytuacji braku jakiegokolwiek wewnętrznego „napięcia” nie miałby on ani dynamiki, ani zdolności sprawczych, które zwykle przypisujemy systemom. System – w sensie użytecznym i ontologicznym – to nie tyle zbiór elementów, ile sieć ich interakcji; bez trwającej wymiany i bez napięć między częściami tracimy systemowość i mamy do czynienia z czystą strukturą.

Wewnętrzny ruch systemu rozumiany jest praktycznie jako wykrywalne zmiany: przepływy energii, wykonanie pracy lub trwałe przestawienie relacji między elementami. Istotne jest tu pojęcie mierzalności – to, co pozostaje poza zasięgiem detekcji, trudno odróżnić od trwałej stałości. Trzeba jednak ostrożnie dodać, że nawet w stanach równowagi, w których siły statycznie się równoważą, nadal zachodzą procesy mikroskopowe (fluktuacje kwantowe) i interakcje; zatem „równowaga” niekoniecznie oznacza dosłowną nieobecność oddziaływań, lecz raczej taki sposób ich znoszenia się, że nie obserwujemy netto zmian makroskopowych. Dopiero absolutny brak zarówno zmian makro-, jak i mikro-skali – ideał teoretyczny – faktycznie unieważniłby mówienie o dynamice systemu.

Wszecławiat musi mieć swoją wewnętrzną energię zdolną do podtrzymania go jako systemu. Zmiany zachodzące na mikropoziomie również potrzebują energii. Dla fizyka, całkowita pula energii układu to suma wszystkich form energii, które znajdują się w obrębie rozważanego układu – kinetycznej, potencjalnej, wewnętrznej (termicznej), pola, masy ($E=mc^2$, Einstein, 1905). W praktyce „całkowita pula energii” to po prostu suma wszystkich form energii, które umówimy się policzyć dla danego obiektu – w laboratorium mierzymy ją przez temperaturę, ciepło i wykonanie pracy; w teorii pola liczymy przez całkę gęstości energii; w kosmologii ciężko mówić o jednej globalnej liczbie, dlatego operuje się raczej na gęstościach i parametrach modelu. Najprościej: przez całkowitą pulę energii układu możemy rozumieć sumę wszystkich form energii, jakie w nim występują – tej związanej z ru-

chem, z oddziaływaniami między jego częściami oraz z samym istnieniem materii; w praktyce nie jest to jedna liczba możliwa do bezpośredniego zmierzenia, lecz pojęcie porządkujące, które pozwala myśleć o układzie jako o całości zachowującej swoją energię w różnych postaciach.

3. RELACYJNA ONTOLOGIA I JEJ KONSEKWENCJE

Rozwiązanie atomistyczne (najmniejsze, niepodzielne byty jako podstawa rzeczywistości) nie wydaje się przekonujące. Owszem, jest ono proste i zdroworozsądkowe. Pojawia się jednak trudne pytanie: jaki jest ontyczny status owych hipotetycznych najmniejszych obiektów? Kolejne pytanie: dlaczego nasze podejście redukcyjne ma polegać na tłumaczeniu tego, co złożone i dynamiczne, przez odwołanie do czegoś, co z definicji ma być proste, statyczne i niepodzielne? Taki sposób myślenia zakłada, że zrozumienie świata polega na rozłożeniu go na części, a następnie na wyjaśnieniu całości przez ich zestawienie. Tymczasem może być dokładnie odwrotnie: to relacje, napięcia i struktury wzajemnych oddziaływań tworzą pozorne „elementy” — a nie elementy tworzą relacje. Być może to, co nazywamy cząstką, jest tylko względnie trwałą konfiguracją pewnego pola procesów, chwilowym ustabilizowaniem się przepływu energii i informacji. Wtedy świat nie byłby mozaiką bytów, lecz splotem wydarzeń i relacji, w których bytowość jest funkcją spójności i trwałości układów, a nie ich niepodzielności.

Nie trzeba zakładać, że elementy najniższego poziomu (mikrocząstki) są ontologicznie podstawowe. Przeciwnie — współczesne przykłady pokazują, że „cząstki” mogą być efektami kolektywnych oddziaływań na dużą skalę: kolektywne wzbudzenia, defekty topologiczne czy quasi-cząstki w układach materii skondensowanej (np. fonony, magnony, pary Coopera w nadprzewodnictwie) zachowują się jak niezależne byty, choć ontologicznie wyłaniają się z całości. To odwraca priorytet: zamiast mikro → makro, mamy często makro → (efektywne) mikro — i to podejście ma poważne konsekwencje dla pojęć takich jak mikrostany, entropia i egzergia.

Jeśli uznamy, że relacje i całościowe wzory oddziaływań są pierwotne, a „cząstki” czy „elementy” wyłaniają się dopiero jako efekty kolektywne, to pojęcie elementu ulega radykalnej rekonstrukcji: nie jest już prostym, atomistycznym bytem o stałej tożsamości, lecz rolą pełnioną w danym poziomie opisu — quasi-bytem, który istnieje wtedy i tylko wtedy, gdy spełnione są określone warunki skali, symetrii i stabilności. Element staje się więc pojęciem relacyjnym i warunkowym: jego granice, trwałość i właściwości zależą od struktury całości, z której się wyłania.

Podobnie „interakcja” przestaje być tylko lokalnym przekazem siły między wcześniej danymi elementami; staje się sposobem, w jaki pola, struktury i globalne uwarunkowania organizują wzbudzenia i przepływy ener-

gii/informacji, skutkując efektywnymi, dyskretnymi egzemplarzami — quasi-cząstkami, defektami topologicznymi czy kolektywnymi modami.

Konsekwencje takiego przesunięcia są dalekosiężne:

- po pierwsze, pojęcia mikrostanu i makrostanu przestają być absolutne — to, co jednoznacznie jest „mikro” w jednym poziomie opisu, może być „makro” wobec jeszcze głębszej, bardziej fundamentalnej struktury.
- po drugie, termodynamiczne i informacyjne właściwości systemu (entropia, egzergia, możliwości pracy) stają się zależne od wyboru skali opisu systemu — od tego, jak szczegółowo ujmujemy jego wewnętrzne zróżnicowanie: to, co wygląda na wzrost entropii z perspektywy jednego poziomu, może być rekonstruowane jako przeniesienie porządku na inną skalę. Entropia (pozorny „chaos”) może być tylko lokalnym zjawiskiem, które w innej skali jawi się jako nowy rodzaj porządku. Można wówczas przyjąć, że w skali całego systemu — wszechświata jako całości — nie sposób mówić o wzroście entropii, bo każda lokalna dezorganizacja jest tylko przesunięciem porządku w inne miejsce; całość zachowuje dynamiczną równowagę, w której porządek i chaos stanowią dwa aspekty tej samej strukturalnej konieczności.
- po trzecie, epistemologicznie i praktycznie: modele powinny dobrać swoje „elementy” nie na podstawie metafizycznej wiary w atomizm, lecz zgodnie z kryterium efektywności opisowej i stabilności przewidywań — inaczej mówiąc, dobry model to ten, który wybiera taki poziom opisu, na którym wyłaniają się trwałe, mierzalne quasi-bity i dla których interakcje mają prostą, użyteczną formę.
- wreszcie, taka ontologia relacyjna otwiera drogę do wyjaśnień, w których porządek i sprawczość są konsekwencją globalnej organizacji, a nie prostą sumą „ruchów” pierwotnych atomów — co ma znaczenie nie tylko dla fizyki, lecz i dla filozofii przyczynowości, teorii systemów i sposobu, w jakim myślimy o zjawiskach emergentnych.

Ujmując wszechświat jako system całościowy, nie jesteśmy zmuszeni do przyjęcia idei śmierci cieplnej jako filozoficznie nieuchronnej. Idea ta wyłania się dopiero po dodaniu konkretnego zestawu założeń — o naturalnej mierze typowości, ergodycznej dynamice (system osiąga wszystkie możliwe stany w długim czasie) i braku mechanizmów odnawiających egzergię. Jeżeli któreś z tych założeń zmienimy — na przykład przyjmujemy, że dynamika jest nieergodyczna (tj. system nie osiąga pełnej równowagi, bo „utknął” w pewnych trajektoriach lub podprzestrzeniach, co pozwala na trwałe nierównowagi i struktury), że istnieją globalne procesy odnawiające dostępne zasoby egzergii, albo że pojęcie „mikrostanu” jest jedynie poziomem efektywnym — to konkluzja o koniecznej utracie zdolności do pracy przestaje być koniecz-

nością ontyczną. Wnioski o przyszłości Wszechświata są zatem warunkowe: śmierć cieplna to ważna hipoteza modelowa, nie przymus metafizyczny.

4. PRZEPIŁYWY ENERGII I INFORMACJI W SYSTEMIE

Przyjmijmy, że system jest obiektem i przestrzenią oddziaływań, zaś jego podsystemy również wchodzą między sobą w interakcje. System oddziałuje na swoje elementy, a elementy oddziałują na system. Oddziaływanie konkretyzuje się jako przepływ energii – w wyniku kolizji czy wymiany jedne części tracą, inne zyskują energię.

Należy jednak rozróżnić dwa przypadki: gdy energia przepływa spontanicznie (od wysokiego do niskiego potencjału), transfer może uwalniać użyteczną pracę; gdy natomiast celem jest ukierunkowane pozyskanie, magazynowanie lub przetworzenie tej energii, często konieczne jest wykonanie pracy sterującej. Ten drugi wymiar – koszt transferu – wiąże się z pokonywaniem pewnej bariery (kinetycznej, potencjałowej lub informacyjnej) oddzielającej obiekty i z wydatkiem energii na kontrolę przepływu. Stąd rozróżnienie między zwykłym przepływem energii a przepływem użytecznym: samo istnienie wymiany nie gwarantuje możliwości wykonywania pracy bez dodatkowych nakładów na sterowanie i magazynowanie.

Skoncentrujmy się na ukierunkowanym przepływie energii. Powiedzmy, że przepływa ona od obiektu A do obiektu B. Który z nich inicjuje ten przepływ? Który z nich ponosi koszt tego transferu? Przy ukierunkowanym przepływie energii nie można jednoznacznie wskazać, kto go inicjuje. W sensie fizycznym przepływ ten wynika z różnicy potencjałów między obiektami: energia przechodzi z miejsca o wyższym potencjale do miejsca o niższym. Jednak w sensie systemowym samo istnienie przepływu zakłada, że obiekt przyjmujący energię wytwarza warunki, które ten przepływ umożliwiają – otwiera przed energią możliwość przejścia, niejako „przyciągając” ją do siebie. Dlatego koszt przepływu ponoszą obie strony: obiekt oddający energię poprzez jej utratę, oraz obiekt przyjmujący – poprzez wysiłek sterowania, porządkowania i utrzymywania warunków odbioru. Można więc powiedzieć, że inicjacja i koszt nie należą do żadnego z obiektów z osobna, lecz do samej relacji, która je łączy. Przepływ energii jest wydarzeniem relacyjnym: obie strony uczestniczą w nim współzależnie, a jego istnienie stanowi cenę utrzymania dynamicznej więzi między nimi. W szerszej perspektywie systemowej każdy taki przepływ jest wyrazem wewnętrznej aktywności całości, która zachowuje swoją spójność właśnie dzięki tym wzajemnym oddziaływanom – utrzymuje się, ponieważ nieustannie się przekształca.

Aby przepływ energii był możliwy i utrzymywał się w określonym kierunku, musi być sterowany – a to sterowanie dokonuje się poprzez informację. Informacja nie jest tu czymś zewnętrznym wobec energii, lecz jej porządku-

jącym aspektem: to sposób, w jaki energia „wie”, dokąd i w jakiej formie ma się przemieścić. Można więc powiedzieć, że informacja nadaje energii wektor, określa jej sens i funkcję w ramach systemu. Im bardziej złożona jest struktura odbierająca energię, tym większy musi być jej zasób informacji pozwalający na rozróżnianie możliwych stanów i kierunków przepływu. Z drugiej strony, samo przetwarzanie informacji wymaga energii — każda operacja porządkująca, każda selekcja czy decyzja pociąga za sobą koszt energetyczny. W ten sposób energia i informacja tworzą obieg wzajemnej zależności: energia umożliwia działanie informacji, a informacja nadaje energii kierunek. System pozostaje sobą dzięki temu sprzężeniu — zachowuje jedność poprzez ciągłe równoważenie między swobodą przepływu energii a porządkiem nadawanym przez informację.

Pobieranie energii można rozumieć jako akt ustanawiania ograniczeń: obiekt A tworzy w przestrzeni B warunki (granice, bariery, kanały), które zawężają możliwy ruch elementów B. W takim ujęciu „odebranie” energii przez A polega nie tylko na przekazie kwantyfikowalnej wielkości fizycznej, lecz przede wszystkim na narzuceniu określonej organizacji ruchu — B traci część swobody, przekazując w zamian energię. To zawężenie przestrzeni stanów B jest jednocześnie nasyceniem go informacją: mniejszy zestaw dopuszczalnych mikrostanów = mniejsza entropia informacyjna. Innymi słowami, informacja tu to praktycznie reguła mówiąca „jak się zachować” — i ta reguła jest wprowadzana przez ograniczenia A. Żeby taki transfer był użyteczny, obiekt pobierający (lub system go obejmujący) musi tę energię przetworzyć i magazynować w formach antycypujących przyszłe potrzeby — np. jako energia potencjalna, wysokouporządkowane wiązania, zapis strukturalny czy stan informacji sterującej. Samo narzucenie ograniczeń wymaga nakładu: budowa i utrzymanie bariery, mechanizmu selekcji czy układu kontrolnego kosztuje energię (i koszt informacyjny — operacje obniżające entropię logiczną mają minimalny koszt; zasada Landauera, 1961 — kasowanie 1 bitu wymaga co najmniej $kT \ln 2$ energii). Z tego powodu każde „złapanie” energii pociąga za sobą nie tylko jej ilościowy transfer, lecz także cenę utrzymania porządku, który temu transferowi służy.

Z tych prostych reguł wyrasta mechanizm rozwoju systemów: konstrukcja i utrzymanie coraz bardziej wyrafinowanych ograniczeń daje lepszą zdolność do kierowania przepływami energii — czyli większą funkcjonalną egzergię — ale równocześnie zwiększa zapotrzebowanie na energię do konserwacji tych ograniczeń. System, który potrafi efektywnie inwestować część pobranej energii w ulepszanie swoich mechanizmów sterowania (informacji), zyskuje przewagę: potrafi wykonywać bardziej skomplikowane funkcje, reprodukować się, magazynować energię i minimalizować straty. To sprzężenie zwrotne — energia \rightarrow ograniczenia/informacja \rightarrow lepsze pozyskiwanie energii — jest rdzeniem procesów wzrostu i ewolucji struktur. Jednak zawsze istnieje kompromis: im więcej porządku (informacji) zakodujesz, tym

więcej energii trzeba poświęcić na jego utrzymanie; jeśli koszty konserwacji przewyższą korzyści energetyczne, struktura staje się nierentowna i ulega rozkładowi lub adaptacji. Tak więc pobieranie energii to tworzenie ograniczeń, ograniczanie ruchu to nasycanie informacją, a użyteczny transfer wymaga zarówno przetwarzania jak i magazynowania tej energii. Z punktu widzenia systemu proces ten tworzy warunek samodoskonalenia — system „uczy się” tworzyć swoje bariery tak, by maksymalizować zdolność do wykonywania pożądanej pracy, przy jednoczesnym balansowaniu kosztów utrzymania porządku.

Gdy obiekt pobiera energię po to, by skuteczniej działać w szerszym systemie, zachodzi równoczesne sprzężenie: energia napływa „z dołu” (od prostszych nośników), podczas gdy informacja spływa „z góry” (reguły i sterowania narzucane elementom). Energia daje zdolność do działania, informacja nadaje temu działaniu kierunek — ogranicza możliwości jednostek, nasycza je informacyjnie i przekształca losowy ruch w pożądane funkcje. To współdziałanie energii i informacji leży u źródła budowy trwalszych, wyższych struktur: porządek powstaje nie tylko dlatego, że coś „ma” energię, lecz dlatego, że całość potrafi tę energię ukierunkować i zmagazynować kosztem utrzymania mechanizmów sterujących.

Energia i informacja nie są tym samym „rodzajem bytu”, ale w praktyce fizycznej da się (w pewnych granicach i za koszty) użyć informacji żeby uzyskać pracę i użyć pracy/energii żeby zapisać/utrzymać informację. Energia i informacja są „powiązane” i — w operacyjnym sensie — wymienne: informacja pozwala zwiększyć wykonalną pracę, a energia jest potrzebna do tworzenia i utrzymania informacji. Informacja (redukcja niepewności) umożliwia wydobycie egzergii i przekształcenie jej w uporządkowaną pracę— a odwrotnie: zapisanie i utrzymanie informacji zawsze wymaga zużycia egzergii i produkcji entropii. Nie oznacza to jednak, że informacja jest „tym samym” co energia: są to różne wielkości (informacja — miara niepewności/korelacji; energia — zasób fizyczny), ale fizyczne procesy łączą je zasadniczo i dają możliwość konwersji z wyraźnymi ograniczeniami (Landauer, 1961; Sagawa–Ueda, 2010 – termodynamika informacji). Z tego powodu mówienie, że „informacja jest fizyczna” ma głęboki sens: zapisywanie, przetwarzanie i używanie informacji zawsze pociąga za sobą wydatki energetyczne i produkcję entropii zgodnie z II zasadą termodynamiki.

5. PODSYSTEM P W KONTEKŚCIE W

Spróbujmy zdefiniować P w ramach krotki $W = (X, \Phi, E, I, \mu, O)$ i rozwinąć jego rolę w wiecznym, relacyjnym wszechświecie. Oto propozycja, jak P mogłoby funkcjonować: Niech $P \subseteq W$ będzie podsystemem, zdefiniowanym jako: $P = (XP, \Phi P, EP, IP, \mu P, OP)$, gdzie:

- $XP \subseteq X$: Podzbiór przestrzeni stanów W , reprezentujący konfiguracje dostępne dla P . Są to stany o wysokim poziomie organizacji (niska entropia lokalna).
- $\Phi P \subseteq \Phi$: Reguły ewolucji specyficzne dla P , które optymalizują przepływy energii i informacji, minimalizując straty egzergii.
- $EP \subseteq E$: Elementy (lub quasi-elementy) P , które mogą być cząstkami, polami, lub emergentnymi strukturami (np. quasi-cząstki, solitony).
- $IP \subseteq I$: Interakcje wewnątrz P i między P a resztą W , które umożliwiają kierowany przepływ energii i informacji.
- $\mu P \subseteq \mu$: Miara ważenia stanów P , preferująca konfiguracje o wysokiej egzergii i niskiej entropii informacyjnej.
- $OP \subseteq O$: Obserwowalne cechy P , takie jak zdolność do pracy, złożoność strukturalna, czy nasycenie informacyjne.

P jest zatem podsystemem, który maksymalizuje lokalną egzergię poprzez sterowanie przepływami energii i informacji, kosztem wzrostu entropii w otoczeniu ($W \setminus P$).

W kontekście konformalnej kosmologii cyklicznej (CCC) (Penrose, *Cycles of Time*, 2010 – wszechświat przechodzi przez nieskończone cykle eonów, z odnowieniem egzergii i dziedziczeniem informacji przez bezmasowe pola), P mogłoby działać w ramach pojedynczego eonu, koncentrując egzergię w trakcie ekspansji, a następnie przekazując informację do następnego eonu poprzez bezmasowe pola (np. fotony, fale grawitacyjne). To pasuje do naszej tezy o dynamicznej równowadze, gdzie lokalny wzrost entropii (w otoczeniu P) jest przesunięciem porządku, a nie globalnym chaosem.

Załóżmy, że P to emergentna struktura (np. supercywilizacja, kosmiczny superorganizm, lub sieć quasi-cząstek), która:

1. Koncentruje egzergię: P aktywnie „wysysa” energię z otoczenia, tworząc gradienty potencjału (np. poprzez grawitację, jak w czarnych dziurach, lub procesy chemiczne/biologiczne w mniejszej skali). To wymaga pracy sterującej, jak opisujemy w części II – np. budowa barier (fizycznych lub informacyjnych) do kierowania przepływów.
2. Przetwarza informację: P zwiększa swoje „nasycenie informacyjne” poprzez zawężanie przestrzeni stanów (XP), co zmniejsza entropię informacyjną. To kosztuje energię (zasada Landauera, 1961), ale pozwala na bardziej efektywną pracę.
3. Przechowuje informację między eonami: W CCC, informacja może przetrwać przejście konformalne (pewne kątowe wzory lub sygnatury, zwłaszcza związane z polami bezmasowymi, z końca wcześniejszego eonu mogłyby przekształcić się w subtelne anomalie w kosmicznym promieniowaniu tła następnego eonu). P mogłoby kodować swoją strukturę w bezmasowych polach, które są niezmiennikami konformalnymi, zapewniając ciągłość w nowym eonie.

Alternatywnie, w modelach wiecznej inflacji (eternal inflation, Guth, *Eternal Inflation and Its Implications*, 2007 – wszechświat składa się z wielu „bąbli”, każdy z innymi prawami fizyki), P mogłoby być strukturą w jednym bąblu, która wpływa na inne bąble poprzez informację (np. kwantowe korelacje). Nasza idea, że entropia jest relatywna do skali opisu, może znajdować szczególny wyraz w kontekście zasady holograficznej ('t Hooft, 1993; Susskind, 1995 – cała informacja w regionie jest zakodowana na jego granicy, np. entropia czarnej dziury \propto pole powierzchni horyzontu), gdzie informacja wszechświata jest kodowana na jego granicy (np. w kosmologii AdS/CFT, Maldacena, 1998 – korespondencja między grawitacją a teorią pola na granicy). W takim ujęciu, globalna entropia W mogłaby być stała, a lokalne zmiany entropii (np. w P) to tylko przesunięcia informacji między skalami.

Nasze ujęcie energii i informacji jako współzależnych jest kluczowe. W P , procesy sterujące przepływem energii (np. tworzenie barier, kanałów) wymagają informacji, a ta wymaga energii. To sprzężenie zwrotne prowadzi do samodoskonalenia: P ewoluuje, by minimalizować straty energii na sterowanie, zwiększając efektywność. W kontekście wiecznego wszechświata, P mogłoby osiągnąć stan „quasi-równowagi”, gdzie koszty sterowania są minimalne, ale nigdy zerowe. To przypomina hipotezę maksymalnej efektywności w biologii (np. organizmy optymalizujące metabolizm) lub w technologii (np. komputery kwantowe minimalizujące straty energii).

6. FORMALNE ROZWINIĘCIE PODSYSTEMU P

Zaproponujemy formalne rozwinięcie naszej krotki W dla podsystemu P , które może być podstawą do dalszych rozważań. Skupmy się na opisie P jako emergentnego systemu sterującego energią i informacją w wiecznym W z niezerową egzergią. Wszechświat W definiujemy jako wieczny, izolowany system z niezerowym poziomem egzergii w stanie naturalnym. Formalnie: $W = (X, \Phi, E, I, \mu, O, \Xi)$, gdzie dodajemy Ξ : globalna egzergia, niezerowa funkcja zachowująca zdolność do pracy w skali całego systemu, $\Xi > 0$.

Podsystem $P \subseteq W$ to emergentna struktura sterująca energią i informacją: $P = (XP, \Phi P, EP, IP, \mu P, OP, I)$, gdzie I : nasycenie informacyjne, miara zawężenia przepływów energii, $I \propto -\Delta S$ (redukcja entropii).

Dynamika P :

- Magazynowanie energii: Energia rozpraszana w $W \setminus P$ jest strukturyzowana w P poprzez I , zapobiegając dyssypacji: $\Delta EP = f(I) > 0$.
- Ograniczanie entropii: Informacja ogranicza mikro/makrostany, zmniejszając energochłonność: $dS/dt \leq 0$ lokalnie w P .
- Samodoskonalenie: ΦP ewoluuje, minimalizując straty: $\min (\Delta E / \Delta I)$.

W CCC, P koduje I w bezmasowych polach, przekazując do następnego eonu.

Dla naszego modelu wszechświata W jako absolutnie izolowanego, wiecznego systemu, zakładamy, że jego stan naturalny obejmuje pewien niezerowy poziom egzergii, co umożliwia ciągłą dynamikę bez konieczności zewnętrznych źródeł. To założenie, inspirowane intuicją, że rozwój wszechświata od Big Bangu (lub jego odpowiednika w modelach cyklicznych) można pojmować jako magazynowanie energii i dyferencjację zawartej w nim informacji, pozwala uniknąć tradycyjnego scenariusza śmierci cieplnej. Zamiast tego, entropia i egzergia stają się relatywne do skali opisu, gdzie lokalny wzrost entropii może być przesunięciem porządku, a informacja pełni rolę strukturyzującą energię, zapobiegając jej pełnej dyssypacji.

Intuicja podpowiada, że rozpraszana energia jest w skali całego wszechświata magazynowana właśnie poprzez nasycanie jej informacją – struktury informacyjne pozwalają na utrzymanie energii w uporządkowanej formie (np. jako energia potencjalna lub kolektywne wzbudzenia), co koresponduje z granicą Bekensteina-Hawkinga (Bekenstein, 1973; Hawking, 1975 – entropia czarnej dziury proporcjonalna do pola powierzchni horyzontu), gdzie entropia regionu jest proporcjonalna do pola powierzchni granicy, ograniczając chaos i umożliwiając emergencję złożoności. Z drugiej strony, nasycanie informacją układów pobierających energię umożliwia zmniejszanie ich energochłonności, jak w systemach złożonych (organizmy, sieci neuronowe), gdzie informacja optymalizuje wykorzystanie energii, ograniczając zarówno mikro- jak i makrostany, co powstrzymuje narastanie entropii.

W tym kontekście, podsystem P powstaje jako emergentna struktura, która koncentruje egzergię W, „wysysając” ją z otoczenia poniżej naturalnego poziomu, co wydaje się możliwe w wiecznym systemie. P musi wykonywać pracę nawet na magazynowanie energii, sterując jej przepływem wewnętrznym poprzez rozbudowane struktury. Mechanizmy ukierunkowujące przepływ (nasycenie informacyjne) wymagają energii, ale pozwalają na przetrwanie bez zewnętrznego pobierania, pod warunkiem zapobiegania wydostawianiu się energii na zewnątrz. P nie kończy rozbudowy, bo praca rozprasa energię (wewnętrznie i zewnętrznie), wymuszając ciągłe doskonalenie i energooszczędność – to sprzężenie zwrotne energii i informacji umożliwia nieskończone funkcjonowanie.

Podsumowując: choć standardowy model kosmologiczny przewiduje nieuchronny spadek egzergii i śmierć cieplną, ekspansja wszechświata napędzana ciemną energią otwiera inne możliwości. Jak wskazał David Layzer (1975), w takim scenariuszu maksymalna możliwa entropia rośnie szybciej niż aktualna – tworzy się trwała „luka entropii”, która podtrzymuje nieskończone gradienty energetyczne i uniemożliwia osiągnięcie równowagi. To bezpośrednio wspiera naszą tezę o naturalnie niezerowym poziomie globalnej egzergii.

Informacja odgrywa tu kluczową rolę – jako wielkość fizyczna zachowana w unitaryjnej ewolucji kwantowej, umożliwia P nie tylko optymalizację procesów sterujących (zgodnie z zasadą Landauera), ale też strukturyzację rozproszonych energii w uporządkowane, kolektywne formy: quasi-cząstki, solitony czy inne trwałe wzbudzenia. W modelach cyklicznych, takich jak konformalna kosmologia Penrose’a, P może kodować swoją strukturę w bezmasowych polach (fotony, fale grawitacyjne), dziedzicząc ją w kolejnych eonach. Podobnie w kosmologiach odbicia w pętlowej grawitacji kwantowej (LQG) – czyli modelach, w których zamiast osobliwości Big Bangu wszechświat „odbija się” od stanu o maksymalnej gęstości, przechodząc z fazy kurczenia się do ekspansji (Ashtekar, 2006; Bojowald, 2001; Rovelli, 2004) – przenosi się energia i informacja z poprzedniego cyklu.

Freeman Dyson (1979) i David Deutsch wyobrażali sobie zaawansowane struktury wykorzystujące ciemną energię do wiecznych obliczeń. Seth Lloyd (2006) widzi wszechświat jako gigantyczny proces przetwarzania informacji, w którym wzrost entropii termodynamicznej współistnieje z eksplozją złożoności. John Wheeler („it from bit”) i Julian Barbour (fizyka bez czasu) wzmacniają relacyjną perspektywę: materia i czasoprzestrzeń mogą być emergentne, a II zasada termodynamiki nie obowiązuje globalnie w nieergodycznym, dynamicznym W.

W ten sposób P nie jest anomalią – to naturalny rezultat holograficznej, informacyjno-energetycznej dynamiki całości. Dzięki nieergodyczności, relacyjności i ciągłemu odnawianiu energii może funkcjonować nieskończenie, nie jako wyjątek od praw fizyki, lecz jako ich wewnętrzna konsekwencja.

7. METAFIZYCZNE SPEKULACJE I IDEALISTYCZNA PERSPEKTYWA

Założyliśmy, że naturalny stan Wszechświata obejmuje pewien niezerowy poziom energii. Poziom ten wystarcza do wykonywania w nim fizykalnie rozumianej pracy, w wyniku której mogą się wyłaniać subsystemy (jak wspomniany P) maksymalizujące lokalną energię poprzez sterowanie przepływami energii i informacji. Jak sugerowaliśmy, proces ten może w zasadzie trwać w nieskończoność, pod warunkiem, że sterująca przepływem energii informacja, którą rozporządza P będzie miała coraz bardziej wysublimowany charakter. Znaczy to, że jest w stanie w sposób coraz bardziej precyzyjny sterować przepływami energii w coraz większej skali – docelowo całym wszechświecie. Ograniczając nieustannie swoją entropię informacyjną może zwiększać zdolność energii do wykonywania pożądanej pracy.

Aby proces ten przebiegał w sposób ukierunkowany, informacja ograniczająca własną entropię musi odnosić się do całości wszechświata. Jądrzem tej informacji musi być zatem pewna teoria wszechświata jako całości (T), w

światle której można go postrzegać jako system, zaś P jako jego subsystem. Powstanie takiej teorii wymaga transcendowania świata, czyli hipotetycznego wykraczania poza dostępną jakimkolwiek podmiotowi jego empiryczną postać. Zakładając, że wszechświat empiryczny ma charakter materialny, wymagająca zewnętrznej perspektywy jego całość musi mieć charakter niematerialny. Uporządkowana sfera niematerialna musi zaś posiadać naturę duchową, czyli autoteliczną (jak osoba).

To otwiera drogę do metafizycznych spekulacji. Najostroźniejszą z nich byłaby taka, że materialny wszechświat nasycy się stopniowo niematerialną informacją. Nieco dalej idącą byłaby taka, że skoro informacja ma naturę niematerialną, a wszechświat zawsze posiadał niezerową egzergię, do której podtrzymania potrzebna jest informacja, to natura wszechświata ma zawsze w jakiejś mierze charakter niematerialny. Teza ta wymagałaby jeszcze dodatkowej przesłanki o bezwzględny ukierunkowaniu informacji (dokładniej: ograniczania entropii informacyjnej) w skali całego wszechświata. Jeszcze dalej sięgałaby teza, że transformacja świata wszechświata polega na zmniejszaniu się roli, jaką odgrywa w nim materia/energia, i zwiększaniu się roli, jaką odgrywa w nim niematerialna informacja.

To jednak dopiero początek. Można bowiem zapytać czy, wobec powyższego, nie byłoby bardziej zrozumiałe przyjąć odwrotny, niż się zwykle zakłada, kierunek ewolucji wszechświata. Skoro jej źródłem jest holistycznie i w gruncie rzeczy niematerialnie pojmowana informacja, to może zamiast przyjmować, że materia nasycy się duchem, należałoby raczej uznać, że to duch nasycy się materią (samo istnienie wszechświata jako wyłącznie materialnego – całkowity brak egzergii i porządku – przypominałoby mitologiczne czy współczesne koncepcje chaosu). To koresponduje z wieloma znanymi z historii filozofii koncepcjami idealistycznymi (np. Platon, Plotyn czy niemiecka filozofia idealistyczna).

W nauce i inspirowanej nią filozofii współczesnej można też spotkać koncepcje nieco ostrożniejsze, które jednak postulują wspomniany odwrócony kierunek ewolucji świata: najpierw systemy, a następnie ich podsystemy czy inne składowe części, które są potrzebne do ich funkcjonowania. Np. Jonathan Schaffer prezentuje stanowisko, zgodnie z którym właściwości i byty będące całościami mają ontologiczne pierwszeństwo przed ich elementami; części istnieją i nabierają sensu przez przynależność do całości. Można je określić jako monizm priorytetowy (Schaffer, *Monism: The Priority of the Whole*, 2010). Uzupełniane byłoby przez stanowisko, które każe patrzeć na ewolucję wszechświata jako hologenezę, czyli proces, w którym w pierwszej kolejności konstytuuje się zintegrowana forma/system, a dopiero w następnej kolejności wydzielają się z niej subsystemy i funkcjonalne składniki. Philip Goff, w artykule „Czy Wszechświat jest świadomym umysłem?” pisze:

współczesna fizyka sugeruje inną, „odgórną” lub „holistyczną” wizję, w której złożone całości są bardziej podstawowe niż ich części. Zgodnie z holi-

zmem stół, który znajduje się przed tobą, nie istnieje na bazie subatomowych cząstek, które go tworzą. To raczej istnienie cząstek subatomowych wywodzi się ze stołu. Ostatecznie wszystko, co istnieje, wywodzi się z podstawowego systemu złożonego: Wszechświata jako całości. (Goff, 2019).

W teorii systemów złożonych mówi się o przyczynowości z góry w dół (downward causation) (Campbell, *Downward Causation*, 1974 – całość determinuje zachowanie części, np. stan organizmu kształtuje komórki). To stwierdzenie jest ściśle powiązane z emergencją i samoorganizacją. W ujęciu, w którym całość ma ontologiczny priorytet, części nie „powstają z niczego”, lecz są generowane i utrwalane przez zestaw wzajemnie wzmacniających się mechanizmów działających na poziomie systemowym. Całość narzuca ogólne ramy działania (tzw. warunki brzegowe) i strukturę powiązań między częściami, które ograniczają możliwe zachowania na poziomie najmniejszych elementów. Jednocześnie stany całej struktury – opisy uproszczone, skupione na głównych cechach (tzw. opisy coarse-grained), kody i reguły przekazywania informacji – pozwalają rozpoznawać lokalne wzorce jako wyraźne, funkcjonujące jednostki. System tworzy zasady i sposoby współpracy, które określają role poszczególnych części, a cele oraz mechanizmy samoregulacji (sprzężenia zwrotne) wybierają i utrwalają te składniki, które najlepiej wspierają działanie całości. W kluczowych momentach rozwoju dochodzi do nagłych zmian stanu (przejścia fazowe) i zrywania równowagi (łamanie symetrii), które powodują skupienie się energii i materii w trwałe, lokalne struktury. Dodatkowo, mechanizmy wsparcia i budowania środowiska (tzw. scaffolding i niche construction) fizycznie i funkcjonalnie podtrzymują rozwój tych struktur. Powtarzalne procesy replikacyjne i funkcjonalna dekompozycja dopracowują i utrwalają wyspecjalizowane podsystemy, podczas gdy topologia sieci (modułowość, centralność) determinuje trwałość ról. Tak złożona sieć mechanizmów pozwala traktować przyczynowość z góry w dół nie jako magiczną konkurencję wobec mikropoziomu, lecz jako kontekstualne ograniczenie i modulację mikrozwisk – daje więc zarówno ontologiczny grunt dla priorytetu całości, jak i realne, testowalne ścieżki wyjaśniające, jak „najpierw systemy, potem części” mogą działać w praktyce.

8. IMPLIKACJE DLA KOSMOLOGII

Nasz model wszechświata jako izolowanego systemu z niezerową egzergią i relacyjną ontologią, gdzie podsystem P steruje przepływami energii i informacji w procesie przyczynowości z góry w dół, niesie znaczące implikacje dla kosmologii, podważając standardowe założenia modelu Λ CDM (standardowy model kosmologiczny z ciemną materią i energią) i otwierając drogę do alternatywnych scenariuszy. Po pierwsze, unikanie śmierci cieplnej dzięki

mechanizmom takim jak „luka entropii” Layzera – gdzie przyspieszająca ekspansja (napędzana ciemną energią) tworzy nieskończone gradienty egzergii – implikuje, że wszechświat może ewoluować wiecznie bez osiągnięcia stanu maksymalnej entropii, co współgra z obserwacjami supernowych typu Ia z 1998 r. i danymi z satelitów Planck. Zamiast tego, P mógłby reprezentować lokalne „oazy porządku”, takie jak galaktyki czy supercywilizacje, które koncentrują egzergię, minimalizując globalny chaos poprzez nasycenie informacyjne, co wspiera hipotezy o holograficznym wszechświecie (granica Bekensteina-Hawkinga), gdzie informacja kodowana na horyzoncie kosmologicznym optymalizuje dynamikę wewnętrzną. Po drugie, odwrotny kierunek ewolucji – od holistycznej całości do materialnych części – kwestionuje Big Bang jako absolutny początek, proponując zamiast singularity „odbicie” w kwantowej grawitacji pętlowej (LQG) (Rovelli, Ashtekar – wszechświat „odbija się” z poprzedniego cyklu zamiast zapadać w osobliwość) lub cykliczne eony w konformalnej kosmologii Penrose’a, gdzie każdy cykl dziedziczy informację (np. poprzez punkty Hawkinga w CMB), umożliwiając ciągłą regenerację egzergii bez zewnętrznych przyczyn. To pasuje do idealistycznej perspektywy, w której niematerialna informacja (duch) nasycza materię, co implikuje, że ciemna energia (ok. 68% wszechświata) mogłaby być manifestacją tej informacyjnej struktury, jak sugerują modele AdS/CFT (Maldacena, 1998), gdzie grawitacja wyłania się z kwantowej teorii pola na granicy. Wreszcie, implikacje obserwacyjne obejmują reinterpretację danych: anizotropie w CMB jako ślady dziedziczonej informacji, a przyspieszenie ekspansji jako mechanizm odnawiania porządku, co mogłoby być testowane przyszłymi misjami jak Euclid czy James Webb Space Telescope. Taki model nie tylko integruje fizykę z metafizyką (np. emanacją Plotyna), ale też zachęca do nowych symulacji kosmologicznych, gdzie wszechświat jako hologeneza stawia całość nad częściami, otwierając drogę do kosmologii idealistycznej, w której świadomość lub informacja jest fundamentalna.

9. ZAKOŃCZENIE

Model Wszczęświata jako systemu izolowanego stanowi próbę przezwyciężenia ograniczeń klasycznych ujęć kosmologicznych i termodynamicznych. Zamiast obrazu świata zmierzającego ku równowadze, proponuje obraz rzeczywistości, która nieustannie się utrzymuje poprzez relacyjne działanie. Niezerowa egzergia, relacyjna natura informacji i dynamiczna równowaga energii oraz entropii składają się na wizję kosmosu, w którym istnienie nie jest stanem, lecz ruchem – ciągłym wysiłkiem zachowania sensu. W tym ujęciu śmierć cieplna Wszczęświata jest jedynie abstrakcyjną granicą, a nie realnym przeznaczeniem: dopóki istnieją relacje, dopóty istnieje działanie, dopóki istnieje działanie – istnieje byt. Wszczęświat trwa, ponieważ nigdy

nie jest skończony w sensie informacyjnym. Każda nowa relacja, każdy akt poznania, każdy lokalny porządek zwiększa jego złożoność i zdolność do trwania. W ten sposób rzeczywistość pozostaje wiecznie żywa – w nieskończonym procesie samoregulacji i samouświadamiania. Być może właśnie to oznacza, że Wszechświat jest „systemem izolowanym”: nie dlatego, że jest zamknięty, lecz dlatego, że cała pełnia działania mieści się w nim samym.

BIBLIOGRAFIA

- A. Ashtekar, *The Big Bounce in Loop Quantum Cosmology*, w: A. Ashtekar (red.), 100 Years of Relativity, World Scientific, 2006, s. 221–242.
- J. D. Bekenstein, *Black Holes and Entropy*, Physical Review D, 7(8), 1973, s. 2333–2346.
- M. Bojowald, *Absence of a Singularity in Loop Quantum Cosmology*, Physical Review Letters, 86(23), 2001, s. 5227–5230.
- D. T. Campbell, *Downward Causation in Hierarchically Organised Biological Systems*, w: F. J. Ayala, T. Dobzhansky (red.), Studies in the Philosophy of Biology, Macmillan, 1974, s. 179–186.
- R. Clausius, *Über die bewegende Kraft der Wärme*, Annalen der Physik, 79(3), 1850, s. 368–397; 79(4), s. 500–524.
- F. Dyson, *Time Without End: Physics and Biology in an Open Universe*, Reviews of Modern Physics, 51(3), 1979, s. 447–460.
- A. Einstein, *Zur Elektrodynamik bewegter Körper*, Annalen der Physik, 322(10), 1905, s. 891–921.
- P. Goff, *Czy Wszechświat jest świadomym umysłem?*, „Filozofuj!”, 19.11.2019.
- A. H. Guth, *Eternal Inflation and Its Implications*, Journal of Physics: Conference Series, 95, 2007, 012004.
- S. W. Hawking, *Particle Creation by Black Holes*, Communications in Mathematical Physics, 43(3), 1975, s. 199–220.
- R. Landauer, *Irreversibility and Heat Generation in the Computing Process*, IBM Journal of Research and Development, 5(3), 1961, s. 183–191.
- D. Layzer, *The Arrow of Time*, Scientific American, 233(6), 1975, s. 56–65.
- S. Lloyd, *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*, Alfred A. Knopf, 2006.
- J. Maldacena, *The Large N Limit of Superconformal Field Theories and Supergravity*, Advances in Theoretical and Mathematical Physics, 2, 1998, s. 231–252.
- R. Penrose, *Cycles of Time: An Extraordinary New View of the Universe*, Bodley Head, 2010.
- C. Rovelli, *Quantum Gravity*, Cambridge University Press, 2004.
- J. Schaffer, *Monism: The Priority of the Whole*, Philosophical Review, 119(1), 2010, s. 31–76.
- L. Susskind, *The World as a Hologram*, Journal of Mathematical Physics, 36(11), 1995, s. 6377–6396.
- G. 't Hooft, *Dimensional Reduction in Quantum Gravity*, arXiv:gr-qc/9310026, 1993.
- E. Verlinde, *On the Origin of Gravity and the Laws of Newton*, arXiv:1001.0785, 2010.

THE UNIVERSE AS AN ISOLATED SYSTEM**ABSTRACT**

This article proposes an original physico-ontological model of the Universe as an absolutely isolated system characterized by non-zero exergy, a relational order, and the ontological primacy of information. Challenging the classical thermodynamic prediction of the heat death of the Universe, it argues that entropy is scale-dependent and that global exergy is continuously sustained by a dynamic coupling of energy and information. Local subsystems ($P \subseteq W$) emerge as self-regulating centers of order, concentrating exergy and reducing informational entropy. The model integrates ideas from conformal cyclic cosmology (Penrose), holographic principles, and downward causation, suggesting a holistic, idealistic cosmology in which the whole precedes the parts (hologenesis). Observational implications include a reinterpretation of CMB anisotropies as imprints of information inherited from previous eons and tests to be performed with Euclid and JWST. The universe is not a system tending toward equilibrium but a self-sustaining relational process—existence as eternal dynamism, meaning, and self-regulation.

Keywords: Universe, ontology, exergy, information, entropy, hologenesis.

O AUTORZE — doktor filozofii, pracował m.in. w Uniwersytecie Ignatianum w Krakowie i ANS w Nowym Targu. Obecnie ANS TWP w Szczecinie. Wydał 6 książek: *Zarys metafizyki absolutności*; *Filozofowanie jako współtworzenie*; *Dociekania. W stronę filozofii niesubstancjalnej*; *Początek i koniec Wszystkiego*; *Historia filozofii jako pamięć i inne szkice*; *Zaczynając od rzeczy pierwszych*. Jego zainteresowania koncentrują się wokół problematyki metafizologii.

Email: stbuda@wp.pl

Aneta Rumak

KONCEPCJA STYLU MYŚLOWEGO LUDWIKA FLECKA A ROZUMIENIE ZJAWISKA WIELOKULTUROWOŚCI

<https://doi.org/10.37240/FiN.2025.13.04>

STRESZCZENIE

Koncepcja stylu myślowego odnajduje swoje zastosowanie w świecie naukowym. Sam autor podaje takie jej zastosowanie, rysując w jej ramach kręgi egzoteryczne a także ezoteryczne i sytuując tam specjalistów w danej dziedzinie oraz laików. Wydaje się jednak, że koncepcję stylu myślowego można z powodzeniem zastosować także w zakresie kultury, szczególnie w odniesieniu do współżycia wielu kultur obok siebie oraz ich wzajemnych relacji. Celem podjęcia rozważań w niniejszym artykule, poza przybliżeniem samej sylwetki mieszkańca Galicji początku XX wieku, jest podkreślenie związków stylu myślowego i zjawiska wielokulturowości. Zestawienie ze sobą stylu myślowego i wielokulturowości, zadanie być może trudne i nowatorskie, wydaje się ważne, gdyż wykazać może niesłabnącą aktualność w dzisiejszym świecie koncepcji Flecka sprzed niemal stu lat.

Keywords: styl myślowy, wielokulturowość, kolektyw myślowy, socjologia wiedzy.

1. WPROWADZENIE

Ludwik Fleck (1896-1961), lwowianin, znany jest przede wszystkim jako mikrobiolog, specjalista z dziedziny bakteriologii. Przykłady do obrazowania koncepcji stylu i kolektywu myślowego czerpał on właśnie z tej dyscypliny, czyli, innymi słowy, ze swej codzienności zawodowej. Najważniejsze dzieło Flecka z zakresu filozofii nauki ukazało się w latach 30. XX wieku, w języku niemieckim, i nie zostało zauważone. Do zwiększonego zainteresowania koncepcjami metodologicznymi i socjologicznymi Ludwika Flecka doprowadziło wydanie *Struktury rewolucji naukowych* Thomasa Kuhna. Amerykański profesor, we wstępie do tej najczęściej cytowanej książki, przyznał się do inspiracji pismami europejskiego autora okresu międzywojnia. Od lat 60. znajomość koncepcji Flecka przebiegała ze zmiennym natężeniem. W Niem-

czach świadomość jego powiązań z nurtem socjologii wiedzy oraz z nazwiskiem Karla Mannheima jest dziś spora. W polskiej myśli socjologicznej, mimo recepcji dzieł lwowianina przed kilku laty, jest on nieobecny a jego koncepcje mało znane. Tymczasem wielce znaczącym wydaje się to, że myślenie jest czynnością społeczną a wiedza żyje w kolektywie. Składa się ona w większej części z tego, co wyuczone niż z tego, co poznane. Poznawanie zaś jest czynnym i żywym nawiązywaniem relacji — jest czynnością zbiorową. Kluczowe dla takich rozważań mogą być słowa Flecka, podkreślające, że jeśli style myślowe są bardzo różne, mogą zachować swoją odrębność w jednym i tym samym indywiduum; jeśli natomiast chodzi o style pokrewne, rozdział taki staje się niemożliwy. Niezgodność stylów myślowych uniemożliwia ich koegzystencję, skazując odnośną osobę na nieproduktywność lub stworzenie szczególnego stylu „strefy pogranicza”. W innym miejscu Fleck dodaje, że style myślowe są nieprzenikalne, pomimo że dopuszczalne jest międzykolektywne krążenie myśli.

Jak jest w istocie, i czy koncepcja stylu myślowego jest pomocna w rozumieniu oraz wyjaśnianiu zjawiska wielokulturowości? Innymi słowy, czy można mówić tu o kompatybilności czy o wykluczaniu się? Znalezienie odpowiedzi na te pytania będzie celem podjęcia przeze mnie rozważań w niniejszym artykule. Teza, dla której utrzymania będzie budowana argumentacja, brzmi: Koncepcja stylu myślowego jest adekwatnym narzędziem do opisu i zrozumienia wielokulturowości. Artykuł składa się z trzech zasadniczych części: omówienia koncepcji stylu i kolektywu myślowego, wyjaśnienia, czym jest wielokulturowość i skąd wynikają trudności z jego definiowaniem oraz zestawienia jej z koncepcjami mieszkańca Galicji z początku XX wieku.

2. ŻYCIORYS NAUKOWY LUDWIKA FLECKA

Prawdopodobnie to czas i miejsce, w jakim przyszło żyć Ludwikowi Fleckowi, sprawiły, że zajął się on medycyną.¹ Mikrobiologiczna strona jego dorobku to swego rodzaju zapis historii tamtych czasów, z kolei zagadnienia serologiczne i mikrobiologiczne, którymi się zajmował, nie pozostały bez wpływu na stworzoną przez niego filozofię nauki. Dorobek z tego zakresu dziś już nie wzbudza szerokiego zainteresowania, również u większości badaczy Flecka.² Dzieje się to z pewnością dlatego, że wiedza w naukach medycznych dezaktualizuje się dość szybko, a współczesnych immunologów nie zainteresują teksty historyczne. Nadto, aby zrozumieć jego niefilozoficzne teksty należy, poza wykształceniem medycznym, posiadać także sporą wiedzę historyczną.

¹ M. Ciesielska, P. Jarnicki, *Ludwik Fleck – mikrobiolog i filozof*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2021, s. 11.

² *Ibidem*, s. 13–14.

Od innych filozofów nauki, gdyż takim mianem można Flecka określać, odróżnia go to, że budował on koncepcje nie na podstawie tekstów innych autorów, lecz na fundamencie własnych doświadczeń. To uważa się za istotną cechę Fleckowskiej twórczości.

Ludwik Fleck urodził się u schyłku XIX wieku (1896 r.) we Lwowie, tam też się kształcił – najpierw w polskim gimnazjum, następnie w Uniwersytecie Jana Kazimierza, na kierunku medycznym. W 1922 roku uzyskał stopień doktora wszechnauk lekarskich. Jakiś czas związany był z macierzystą uczelnią, gdzie pracował jako asystent Rudolfa Weigla w katedrze Biologii Ogólnej. Specjalizację z bakteriologii uzyskał w Wiedniu, w Instytucie Seroterapeutycznym u prof. Krausa. Po powrocie do Lwowa kierował, w latach 1928-1935 laboratorium bakteriologicznym i chemicznym w Zakładzie Ubezpieczeń Społecznych. W czasie sowieckiej okupacji tego miasta pracował w tzw. Wydziale Medycznym, wydzielonym z uniwersytetu, jako adiunkt w Katedrze Bakteriologii. Z kolei niemiecka okupacja miasta, od czerwca 1941 roku, doprowadziła do tego, że wraz z innymi żydowskimi uczonymi Fleck rozpoczął pracę w Instytucie nad Tyfusem Plamistym i Wirusami prof. Rudolfa Weigla. Praca ta chroniła dłuższy czas przed represjami ze strony okupanta. W styczniu 1943 roku żydowscy uczeni z Instytutu kierowanego przez Rudolfa Weigla zostali wywiezieni, najpierw do Auschwitz, a następnie do Buchenwald. Fleck był przez Niemców postrzegany jako specjalista w dziedzinie zwalczania tyfusu, dlatego uniknął śmierci i przetrwał wojnę. Podobny los spotkał żonę i syna. Po zakończeniu wojny cała rodzina trafiła do Lublina, gdzie Ludwik Fleck objął kierownictwo Zakładu Mikrobiologii Medycznej UMCS, a potem Akademii Medycznej w Lublinie. Druga połowa lat czterdziestych to habilitacja, docentura, tytuł profesora nadzwyczajnego i wreszcie, w 1950, tytuł profesora zwyczajnego. Od 1952 roku kierował Zakładem Bakteriologii i Immunologii w Instytucie Matki i Dziecka w Warszawie. Był członkiem Polskiej Akademii Nauk oraz laureatem licznych nagród i odznaczeń państwowych. W 1957 roku, w ślad za jedynym synem, wyjechał wraz z żoną do Izraela, gdzie, w Instytucie Bakteriologicznym w Nes-Syjon, kontynuował badania nad zjawiskiem leukergii. Fleck kierował także Departamentem Patologii Eksperymentalnej oraz Izraelskim Instytutem Badań Biologicznych, prowadząc badania nad mechanizmami odpornościowymi. Na jego koncie jest także uczestnictwo w licznych międzynarodowych kongresach oraz prowadzenie wykładów w Brazylii, we Francji, w USA i w ZSRR.

Zainteresowania badawcze Flecka koncentrowały się wokół problematyki bakteriologicznej, ze szczególnym uwzględnieniem diagnostyki chorób zakaźnych i serologii. Pierwszą znaczącą pracę opublikował w 1931 roku (współautorstwo z I. Heschelesem), w której dowodził, że wyciągi z *Proteus X 19* podane podskórnie osobom zdrowym dają w efekcie miejscowy odczyn zapalny. Natomiast te same wyciągi aplikowane chorym lub ozdrowieńcom

nie skutkowały odczynem zapalnym.³ Zauważył tym samym, że osoby szczepione przeciw durowi plamistemu wykazują zmniejszoną wrażliwość na wywołanie tegoż odczynu zapalnego (Fleck zaobserwował go już w 1930 roku, nazywając egzantynowym).⁴

Jeśli idzie o nauki medyczne oraz przyrodnicze, Fleck, wraz z E. Altenbergiem, ogłosił w 1931 roku wyniki swoich badań nad rozmieszczeniem leukocytów we krwi, oparte na zastosowaniu rachunku prawdopodobieństwa. Badania te doprowadziły do odkrycia zjawiska leukergii, nieznanego dotąd, czyli zjawiska zwiększonej lepkości i skłonności do aglomeracji białych ciałek. Fleck łączył to zjawisko z wczesną reakcją obronną organizmu, która występuje już w serologicznym okresie zakażenia. Zjawisko leukergii zostało odkryte w czasie wojny, w 1942 roku, jednak w wojennych warunkach trudno było o konieczne badania eksperymentalne, dlatego też prace na ten temat ukazały się dopiero w 1946 i 1947 roku.

Prace filozoficzne, początkowo niedostrzegane, po latach przyniosły autorowi światowy rozgłos. Pierwszą ważną pracą Flecka z zakresu filozofii medycyny był artykuł *O niektórych swoistych cechach myślenia lekarskiego* (1927). Przewodnim było tam założenie, że o ile w naukach przyrodniczych uwaga uczonego skupia się na cechach i zjawiskach prawidłowych, o tyle w medycynie myślenie ukierunkowuje się na zmiany patologiczne. Autor starał się więc udowodnić, że uogólnione pojęcie choroby jest z gruntu fałszywe i niepotrzebne.⁵ Można mówić co najwyżej o istnieniu zjawisk patologicznych, najczęściej rozpatrywanych w każdym indywidualnym przypadku. Znaczy to, że Fleck po raz pierwszy zaprzeczył obiektywnemu istnieniu faktów lekarskich. W kolejnej pracy, *Zur Krise der „Wirklichkeit“*. *Die Naturwissenschaften* dowodził, że rzeczywistość nie ma charakteru obiektywnego, jest za to ściśle uzależniona od indywidualnego doświadczenia. Czynny i świadomy akt poznawczy jest uzależniony od stopnia wykształcenia, obciążenia tradycją i predyspozycji psychicznych.

Obie wspomniane publikacje były swoistym preludium do dzieła głównego, wydanego w 1935 roku *Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv*. Ani jednak ta książka, ani artykuły pisane po polsku czy po niemiecku nie wzbudziły zainteresowania. Fleck stworzył tymczasem nową, antypozytywistyczną, przeczącą obiektywności naszego poznania filozofię,

³ <https://gigancinauki.pl/gn/biogramy/82990,Fleck-Ludwik.html#> [dostęp: 20 listopada 2024r.].

⁴ Badania moczu chorych na dur plamisty prowadził Fleck najpierw w getcie lwowskim, następnie zaś w obozach w Auschwitz oraz w Buchenwaldzie. Celem tych badań było poszukiwanie swoistych substancji antygenowych, które, jeśli tylko byłyby obecne, mogły pełnić funkcję wskaźnika umożliwiającego wykrycie choroby. Występowanie substancji antygenowych w moczu chorych na dur plamisty otwierało drogę do uroserologicznego rozpoznania duru. Stwarzało to nadzieję, że moc mógłby posłużyć do sporządzenia prostej i taniej szczepionki, zob. ibidem.

⁵ R. W. Gryglewski, *Ludwik Fleck* [w:] *Giganci nauki. Biogramy*: <https://gigancinauki.pl/gn/biogramy/82990,Fleck-Ludwik.html#> [dostęp: 20 listopada 2024r.].

opierając ją w zasadzie na dwóch pojęciach: stylu myślowego oraz kolektywu myślowego. Poznanie ma charakter subiektywny, zależny od uwarunkowań psychicznych i socjologicznych. Formułowanie sądów i koniecznych w nauce uogólnień (definicji) odbywa się na podstawie obowiązujących w danej społeczności przekonań i przyjętych zasad, na drodze swoistego kompromisu i przenikania się stylów myślowych, co Fleck nazywa „kolektywem myślowym”.⁶ Koncepcje te stały się inspiracją dla Thomasa Kuhna, które we wstępie do *Struktury rewolucji naukowych* z 1962 roku przywołał nazwisko lwowianina. Publikacja książki Kuhna rozpoczęła proces przywracania do robku Flecka pamięci ludzkiej, którego punktem kulminacyjnym wydaje się wydanie angielskiego przekładu *Powstania i rozwoju faktu naukowego*. Od lat 80. systematycznie wzrasta zainteresowanie pismami filozoficznymi Ludwika Flecka.

3. LUDWIKA FLECKA TEORIA POZNANIA. KONCEPCJA STYLU I KOLEKTYWU MYŚLOWEGO

O stylu myślowym pisze Fleck pisze, że składa się on „jak każdy styl, z pewnego określonego nastroju i realizującego ten nastrój wykonania. Nastrój posiada dwie ściśle ze sobą związane strony – gotowość do selektywnego odczuwania i do odpowiednio ukierunkowanego działania. Stwarza adekwatne dla siebie wyrazy: religia, nauka, sztuka, obyczaj, wojna itp., stosownie do przewagi pewnych kolektywnych motywów i kolektywnie zastosowanych środków. Możemy więc definiować styl myślowy jako ukierunkowane przestrzeganie wraz z odpowiednią obróbką myślową i rzeczową tego, co postrzegane [...]. Charakteryzują go wspólne cechy problemów, którymi kolektyw jest zainteresowany: sądów, które uważa za oczywiste; metod, których używa się jako środków poznawczych”.⁷

Nośnikiem stylu myślowego jest kolektyw myślowy.⁸ Może się on wytworzyć nawet wtedy, gdy tylko dwie osoby wymieniają swe myśli – już wtedy wytwarza się specyficzny nastrój, który znika, gdy uczestnicy rozstaną się, i powracający, gdy ponownie dochodzi do ich spotkania. Każdy z nas należy do wielu kolektywów myślowych równocześnie, najczęściej nie zdając sobie z tego sprawy. W obrębie kolektywów myślowych funkcjonują pewne pojęcia, aksjologie, przesady, co do których zasadności i realności utwierdzają się ich członkowie w wyniku stałej wymiany myśli. Wspólny styl myślowy, który łączy członków wspólnoty, wytwarza też pewnego rodzaju przymus czy ra-

⁶ Ibidem.

⁷ L. Fleck, *Powstanie i rozwój faktu naukowego w: Psychosocjologia poznania naukowego. Powstanie i rozwój faktu naukowego oraz inne pisma z filozofii prawa*, Cackowski, Z., Symotiuk, S. (red.), Lublin 2006, s. 120–121.

⁸ Tamże, s. 68.

czej stałą gotowość do określonego myślenia i działania. Od stylu myślowego zależy fakt naukowy – u Flecka konstruowany, a mimo to obiektywny. Istnieją kolektywy myślowe chwilowe, aktualne jedynie czasowo, oraz kolektywy stałe, z własnymi strukturami i z niejednokrotnie długą tradycją. W przypadku tych ostatnich Fleck podkreśla ich charakterystyczną dwudzielność: z jednej strony widoczny jest krąg ezoteryczny, tworzony przez specjalistów, z drugiej – krąg egzoteryczny, w pewien sposób zewnętrzny wobec właściwej struktury myślowej. Krąg ezoteryczny dysponuje wiedzą fachową, z kolei krąg egzoteryczny posiada wiedzę popularną (zwolennicy, sympatycy, opinia publiczna).⁹ Między kolektywami może dochodzić do wymiany myśli, przy czym każdorazowa taka wymiana oznacza przesunięcie w zakresie znaczenia pojęć i wartości myśli, co w konsekwencji może doprowadzić do zmiany postrzegania i do stworzenia nowych faktów.

Warunkiem istnienia kolektywu myślowego jest „wymiana myśli”, ukierunkowana przez „nastrój myślowy”. Długo działający nastrój myślowy doprowadził do wyizolowania się kolektywu naukowego, kolektywów religijnych, artystycznych. Od izolacji formalnej ważniejsza jest izolacja treściowa kolektywu (styl) jako szczególnego „świata myśli”.¹⁰ Aby stać się członkiem jakiegoś kolektywu, trzeba przejść okres terminowania, podczas którego zachodzi autorytatywna „sugestia myślowa”, czyli wdrożenie w dany styl myślowy, w „zamknięty styl myśli”. Kolektyw to nie prosta suma jednostek, lecz nowa jakość, która określa nie tylko myślenie indywidualne, ile właśnie styl myślowy, którego poszczególne jednostki są nosicielami oraz nastrój myślowy właściwy dla danego kolektywu (nastrój odpowiada za kierunek, w jakim dany styl myślowy się rozwija).

Swoje rozważania teoriopoznawcze Fleck rozpoczyna od medycyny. Wiedzę tam zdobytą przenosi na nauki przyrodnicze i konstruuje ogólną teorię poznawania, tj. porównawczą socjologię myślenia, która kolektyw myślowy i styl myślowy czyni kategorią nadrzędną, zdając sobie najpewniej sprawę z twórczej roli kultury w procesie poznania.¹¹ I chociaż jako czas narodzin projektu socjologii wiedzy naukowej uznaje się połowę lat 70. XX wieku, to za jego prekursora Rydlewski proponuje uważać Flecka właśnie.¹² Oczywiście, wcześniejsze były dzieła fundamentalne dla tej dyscypliny: *Problem socjologii wiedzy* Maxa Shelera oraz *Ideologia i utopia* Karla Mannheima; Fleck

⁹ Należy mieć tu na myśli przede wszystkim kolektywy naukowe; je bowiem, w odróżnieniu od kolektywów religijnych, cechuje wielość centrów ezoterycznych; por. P. Jarnicki, *Ludwika Flecka nauka bez prawdy?*, Przegląd Filozoficzny – Nowa Seria, 2, 2010, s. 65.

¹⁰ P. Jarnicki, *Antytotitarne motywy twórczości Ludwika Flecka w: Nauka w kontekście wzorców kultury*, Płonka-Syroka, B., Kaźmierczak, A. (red.), DiG, Warszawa 2011, s. 94.

¹¹ Fleck wskazuje na biologię jako tę dziedzinę, która nauczyła go myśleć historycznie. Konstrukttywizm w sposobie percepcji odkryć naukowych pojawił się u Flecka między opublikowaniem w 1927 roku tekstu *O niektórych swoistych cechach myślenia lekarskiego* a napisaniem w roku 1929 *O kryzysie „rzeczywistości”*.

¹² M. Rydlewski, *W stronę teorii i historii kultury. Spojrzenie na teorię stylów myślowych Ludwika Flecka*, Etnografia Polska, 1–2, 2009, s. 116–117.

jednak dostrzegają i był świadomy pewnej niekonsekwencji. Szło o utożsamianie społecznych uwarunkowań wiedzy z deformacją poznania. W związku z tym możliwa byłaby wiedza nieuwarunkowana społecznie, czyli wolna od wpływu kultury. Z kolei założenie to jest sprzeczne z podstawową tezą socjologii wiedzy, głoszącą, że wiedza jest uwarunkowana społecznie. Sojak zauważa, że niekonsekwencja ta wynika nie tyle z braku świadomości problemu, co z chęci ucieczki przed zarzutem relatywizmu.¹³ Fleck w tworzeniu koncepcji stylu myślowego zdaje się rozumieć tę niekonsekwencję i szukać metody wyjścia z impasu. Zarzucano Fleckowi, że z jego teorii poznawania jako nauki o historycznym i socjologicznym rozwoju stylów myślowych „wynika niewątpliwy relatywizm poznawczy, który przyznaje różnym obrazom rzeczywistości, wyrosłym z różnych stylów myślowych, jednakową ważność, może nawet wtedy, gdy te obrazy są sprzeczne”.¹⁴ Na ten zarzut Fleck odpowiada słowami: „Z teorii stylów myślowych nie wynika żaden relatywizm poznawczy. »Prawda« jako aktualny etap przemian stylu myślowego jest zawsze tylko jedna: jest ona stylowo bez reszty zdeterminowana”.¹⁵ Zdaniem Bogdana Dziobkowskiego. Fleck jest realistą wewnętrznym.¹⁶ Koncepcja Hilary’ego Putnama, nazywana także realizmem pragmatycznym, „sprowadza się do stanowczego twierdzenia, że realizm *nie* wyklucza względności pojęciowej. Można być zarazem realistą i relatywistą pojęciowym”.¹⁷ Według realisty pojęciowego pojęcia nie mają „absolutnych” znaczeń, ale wiele różnych zastosowań.

W koncepcji Flecka odnaleźć można tzw. „zjawisko równoważności opisów”. Mnogość przykładów z historii nauki opisanych przez niego wskazuje na istnienie wielu niewspółmiernych, nieprzekładalnych na siebie stylów myślowych. O prawdzie pisze Fleck, że jest to „zgodne ze stylem, jedyne możliwe rozwiązanie”. Prawda nie jest względna, ani subiektywna w popularnym tego słowa znaczeniu. Prawda jest zawsze, lub prawie zawsze, w pełni zdeterminowana w obrębie jakiegoś stylu myślowego. Fleck zdaje się wykluczać, że to samo zdanie jest dla jednej osoby prawdziwe, dla innej zaś fałszywe.¹⁸ Wyjaśnia, że jeśli obie są członkami tego samego kolektywu my-

¹³ R. Sojak, *Czy spór o relatywizm można rozwiązać empirycznie?* w: *Teoretyczne podstawy socjologii wiedzy*, t. 1, Bytniewski, P., Chałubiński, M. (red.), UMCS, Lublin 2006, s. 287-302.

¹⁴ T. Bilikiewicz, *Uwagi nad artykułem Ludwika Flecka »Nauka a środowisko«* w: Fleck, L. *Powstanie i rozwój faktu naukowego*, UMCS, Lublin 1986, 189-192.

¹⁵ L. Fleck, *Powstanie i rozwój faktu naukowego. Wprowadzenie do nauki o stylu myślowym i kolektywie myślowym*, przeł. M. Tuskiewicz, Wydawnictwo Lubelskie, Lublin 1986, s. 184; W. Sady, *Fleck. O społecznej naturze poznania*, Prószyński i S-ka, Warszawa 2000, s. 79-83.

¹⁶ B. Dziobkowski, *Realizm wewnętrzny a relatywizm*, Przegląd Filozoficzny - Nowa Seria, 3, 2006, s. 29.

¹⁷ H. Putnam, *Wiele twarzy realizmu i inne eseje*, przeł. A. Grobler, PWN, Warszawa 1998, s. 342.

¹⁸ Zdanie to może rodzić wątpliwości, jeśli weźmie się od uwagę polemikę Flecka z Tadeuszem Bilikiewiczem. Tam właśnie ma miejsce dyskusja na temat „normalnej ręki, która ma pięć palców”. Bilikiewicz sądził, że jest czymś niemożliwym, by powstały choćby dwie niepokrywające się ze sobą teorie na temat takich faktów, jak: „normalna ręka ma pięć palców”, nawet jeśli badacze będą należeć do najodleglejszych kolektywów myślowych. Zdaniem Bilikiewicza nie jest to kwestia stylu, ale oglądu. Fleck wyjaśniał, że to, co rozumiemy przez „normalną rękę” uzależnione jest od stylu. Co

slowego, to określone zdanie jest dla obu albo prawdziwe, albo fałszywe. Jeśli nie należą do tych samych kolektywów, to nie mamy do czynienia z tym samym zdaniem.

Koncepcje stylu i kolektywu myślowego nie obywają się bez zarzutów kierowanych w ich stronę. Wysunięto ich już co najmniej kilkanaście, rozmiary niniejszego artykułu pozwalają na przywołanie jedynie kilku najważniejszych. Pierwszy z nich pojawia się już przy okazji potrzeby wytyczenia granicy między poszczególnymi stylami i kolektywami myślowymi. Jeśli tylko spytamy o kryterium grupowego zróżnicowania myślowego ludzi, to u Flecka znajdziemy wyjaśnienie, że kolektywem myślowym jest „wspólnota ludzi związana wymianą myśli lub wzajemnym oddziaływaniem intelektualnym”. Jeśli zaś skupić się na kwestii oddziaływania intelektualnego, można dojść do wniosku, że kolektywy myślowe polegają na społecznym zróżnicowaniu ludzi możliwym do stwierdzenia w oparciu o to, czy mogą się zrozumieć, czy nie.¹⁹ Pojęcie rozumienia jest jednak problematyczne, o tyle problematyczne, że jego proces odbywa się niejako wewnątrz człowieka. Jest trudny do zaobserwowania, a tego typu próba przyporządkowania ludzi do kolektywów myślowych musiałaby prowadzić do empirycznego stwierdzenia, kto co jak rozumiał. Niewykluczone, że tego rodzaju dylematy skłoniły Flecka do zachowania dystansu wobec pojęcia rozumienia i stwierdzenia, że „dopiero socjologia myślenia może nam wytłumaczyć zagadnienie porozumienia i nieporozumienia się ludzi”.²⁰ Fleckowi idzie więc najpewniej o porozumienie, nie rozumienie.

Kolejny problem można zaobserwować przy pytaniu o przyporządkowanie ludzi do kolektywów myślowych. Dla Flecka pierwszym i zasadniczym zjawiskiem, jakie powinna badać nauka o poznawaniu, jest „grupowe zróżnicowanie myślowe ludzi”. Pojawia się też „członek” na określenie przynależności do danego kolektywu myślowego. Słuszne wydaje się pytanie o źródło takiej przynależności. Z punktu widzenia socjologii jedynie organizacje mogą mieć członków. U Flecka taka, nazwijmy ją oficjalną, wspólnota, jak chociażby religijna, nie pokrywa się całkowicie z kolektywem myślowym. Fleckowi idzie bardziej o nieformalne pojęcie członkostwa. W tym rozumieniu na przykład jakiś naukowiec nienależący do swojej rodzimej organizacji na-

innego to wyrażenie będzie oznaczało dla lekarza, co innego dla laika. „Całe zaś zdanie, w naszym stylu niewątpliwy fakt, może dla innych stylów być bez sensu” – pisał Fleck w odpowiedzi na polemikę Bilikiewicza. Odnosił się przy tym do licznych prymitywnych społeczeństw, które tylko dla liczb 1,2,3 mają nazwy odrębne. Papuasi używają na określenie liczby pięć słowa: „ręka”, dziesięć – „dwie ręce”. Lecz Fleck mówi o dwóch teoriach nie o jako sprzecznych, a niewspółmiernych; por. L. Fleck, *Odpowiedź na uwagi Tadeusza Bilikiewicza* w: Fleck L., *Psychosocjologia poznania naukowego*, op. cit., s. 287-288.

¹⁹ T. Möller, *Uwagi krytyczne na temat pojęć kolektyw myślowy, styl myślowy i wymiana myśli*, w: *Style myślowe i fakty. Artykuły i świadectwa*, Werner, S., Zittl, C., Schmaltz, F. (red.), Instytut Filozofii i Socjologii Polskiej Akademii Nauk, Warszawa 2007, s. 307.

²⁰ L. Fleck, *Powstanie i rozwój faktu naukowego* [w:] *Psychosocjologia poznania naukowego*, op. cit., s. 32.

ukowej mógłby być, w szerokim sensie tego pojęcia, członkiem społeczności naukowej. Taki rodzaj członkostwa wynikałby z faktu odbytych studiów, ich ukończenia i wymiany myśli z innymi naukowcami o tym samym swoistym stylu myślowym.

Reasumując, rozgraniczenie pomiędzy różnymi kolektywami myślowymi jest u Flecka niejasne; przynależność ludzi do kolektywów myślowych – problematyczna, a pojęcia używane przez tego badacza pochodzą z tradycji świadomości, której nie sposób odnieść do teorii komunikacji. Można ocenić, że Fleck przygotował koncepcję nie do końca dojrzałą, nacechowaną niespójnością. Jednak inne partie dzieła Flecka przekonują, że do teorii komunikacji filozof odniósł się z należytą starannością. Świadczą o tym części, w których Fleck przedstawia analizy dokonywane na konkretnych przykładach i w których mniejszy nacisk kładzie na zaprezentowanie grupowego zróżnicowania myślowego ludzi, a większy na porównanie źródeł pisanych i zawartych w nich stylów myślowych.

4. WIELOZNACZNOŚĆ WIELOKULTUROWOŚCI. PROBLEMY DEFINICYJNE

O wielokulturowości można mówić jako o pewnym fenomenie. Nieodzwonnie dziś towarzyszyć może wrażenie, że jest to modne hasło, odnoszące coraz większą popularność w zglobalizowanym świecie, w którym wielokulturowość jest albo wysuwana na czoło wszystkich tych zjawisk, lub jest tłumiona, gdyż globalizacja dąży do jednolitości w każdej sferze. Tymczasem wielość kultur nie wynika tylko i wyłącznie z wieloetniczności czy wielonarodowości. Każde społeczeństwo jest wielokulturowe, jeśli przyjąć, że dycho-
tomiczność podziału: miasto/wieś może już na tę wielokulturowość wskazywać. Dlatego pojęcie wielokulturowości jest bardzo trudno zdefiniować jednoznacznie. Trudności definicyjne odsłaniają się jeszcze bardziej, gdy głębsza refleksja dotyczyć będzie problemów społeczeństw nazywanych wielokulturowymi, w tym kwestii wymieszania się kultur, co obrazuje metafora „tygla”, lub też w kwestii bycia kultur obok siebie, bez nawiązywania bliższych relacji a przy zachowaniu dystansu i tożsamości, co z kolei oddawać ma metafora „półmiska z sałatkami”. Wielokulturowość bywa odróżniana od multikulturalizmu jako od polityki w zarządzaniu społeczeństwem „wielu kultur”.

Można zaobserwować niemałe zamieszanie z rozumieniem wielokulturowości w literaturze. Jest wielokulturowość bowiem i ideą, i ideologią, jest i faktem społecznym, ale jest i polityką. Bywa wymieniana zamiennie z multikulturalizmem, bywa także od niego odróżniana.

Podmiotowo zagadnienie wielokulturowości dotyczy różnorodnych grup etnicznych. Will Kymlicka, dokonawszy kompleksowej analizy problemu,

zaliczył do niego nie tylko mniejszości narodowe, lecz także i imigrantów (ekonomicznych i politycznych, dysponujących możliwością nabycia obywatelstwa), migrantów nielegalnych i czasowych (nie traktowanych jako obywatele), członków izolacjonistycznych grup etniczno-religijnych (tzw. „obywateli częściowych”) oraz Afroamerykanów, którzy, żyjąc w Stanach Zjednoczonych równie długo jak biała ludność, stanowią oddzielną kategorię społeczną.²¹ Jest tak pomimo tego, że długo nie mogli oni otrzymać obywatelstwa, wynikających z niego uprawnień, jednocześnie nie posiadając jednolitej kultury i języka i nie będąc tym samym grupą homogeniczną.

Rozumienie terminu „wielokulturowość” opiera się na klasycznym ujęciu biorącym za podstawę etniczność i utożsamia wielokulturowość z zetknięciem się ze sobą przedstawicieli przynajmniej dwóch narodowości.²² Początkowo to właśnie etniczność była elementem określającym kulturę i była związana z zajmowanym obszarem geograficznym. Wielokulturowość rozumiano jako zderzenie kilku różnych etniczności, często w wyniku ekspansji terytorialnej. Następnie zaczęto obejmować tym terminem całość stosunków społecznych i kulturowych we współczesnych społeczeństwach, włączając w definicję wielokulturowości także cechy nie związane z przestrzennym aspektem funkcjonowania grupy. Jako przykład podejścia widzącego zderzenie kultur w czasie Hanna Mamzer podaje koncepcję społeczności postfiguratywnych, kofiguratywnych i prefiguratywnych pochodzącą od Margaret Mead.²³ Wielokulturowość w tym aspekcie jest spotkaniem kultur powstałych w różnym czasie, chociaż na tej samej przestrzeni. Swoistą jednostką miary jest w tym przypadku jedno pokolenie. Mamzer, chcąc zaznaczyć trudności z definiowaniem wielokulturowości, wymienia cztery możliwe podejścia do tego zjawiska. Jej zdaniem, wielokulturowość jest albo zbiorem zasad i procesem realizacji współżycia społecznego w warunkach realizacji pluralizmu etnicznego i kulturowego, przy wyrównywaniu szans uczestnictwa (to raczej postulat, jak powinno spotkanie kultur wyglądać), albo jest ona spotkaniem grup o różnych programach, które określa się w oparciu o dominujące w danej grupie wartości.²⁴ W trzecim rozumieniu, wielokulturowość to każde spotkanie kultur, które ma interakcyjny charakter. Owa interakcyjność wymusza nawet, by wielokulturowość zastąpić in-

²¹ W. Kymlicka, *Współczesna filozofia polityczna*, przeł. A. Pawelec, Aletheia, Warszawa 2009, s. 420-425.

²² H. Mamzer, *Tożsamość w podróży. Wielokulturowość a kształtowanie tożsamości jednostki*, UAM, Poznań 2002, s. 32.

²³ M. Mead, *Kultura i tożsamość. Studium dystansu międzykulturowego*, przeł. J. Hołówka, Warszawa 2000; H. Mamzer, *Tożsamość w podróży...op. cit.*, s. 33.

²⁴ Wartości, które można wymienić, to: równouprawnienie (prym wiedzy w kulturze feministycznej); dążenie do obrony pokoju (kultura pacyfistyczna); różnice rasowe (kultura rasistowska). Poszczególni członkowie tak wyróżnionych zbiorowości nie muszą się czuć do przynależności do niej; do grupy zostają zaliczeni na podstawie jednej cechy (to subkultury, których spotkanie nazywane jest „wielokulturowością”), por. H. Mamzer, *Tożsamość w podróży...op. cit.*, s. 34.

nym terminem: międzykulturowość.²⁵ I wreszcie, wielokulturowość to współwystępowanie na tej samej przestrzeni (albo w bezpośrednim sąsiedztwie bez wyraźnego rozgraniczenia, albo w sytuacji aspiracji do zajęcia tej samej przestrzeni) dwóch lub więcej grup społecznych o odmiennych cechach dystynktywnych: wyglądzie zewnętrznym, języku, wyznaniu religijnym, układzie wartości, które przyczyniają się do wzajemnego postrzegania odmienności z różnymi tego skutkami. Stosunki, jakie panują pomiędzy przedstawicielami tych różnych kultur, można ująć w kilka zasadniczych kategorii: otwarty antagonizm, pasywny antagonizm, segregacja jawna lub skrywana, koegzystencja pozorna polegająca na wzajemnej akomodacji, asymilacja połączona z pełną wzajemną akceptacją i współdziałaniem.²⁶ Innego podziału wielokulturowości dokonuje Marta Kubiszyn, akcentując klasyfikację opartą na wydzieleniu trzech rodzajów zjawiska. W rozumieniu wąskim Kubiszyn traktuje społeczeństwo jako jeden organizm – to w jego obrębie dostrzec można istnienie odmiennych kultur bądź wyznań. W rozumieniu nieco szerszym autorka nawiązuje do relacji, jakie zachodzą między odmiennymi kulturami. Najszersze rozumienie dotyczy przestrzennego przekraczania granic kulturowych oraz powstawania nowych nieznanymi wcześniej struktur.²⁷

Różnorodne rozumienie i podziały wielokulturowości pokazują, że współcześnie stanowi ono jedno z największych wyzwań, jakie stoją przed człowiekiem. „Inność” jest zagadkowa, dziwna, może napawać lękiem przed nieznanym. W tej części artykułu podałam kilka definicji wielokulturowości, nie skupiając się jednocześnie na tych, które odnoszą się do wielokulturowości jako polityki, a wyróżniając bardziej te, które zjawisko wielu kultur postrzegają interakcyjnie i na ten właśnie aspekt kładą nacisk. W mojej ocenie takie rozumienie będzie szczególnie przydatne w analizowaniu związku wielokulturowości i stylu myślowego Ludwika Flecka. Analiza ma za zadanie ocenić

²⁵ Jak pisze Leszek Korpowicz, międzykulturowość „lepiej oddaje i akcentuje transkulturowy charakter procesów wzajemnego uczenia się, włączania w obszar własnych standardów wartości kulturowych innych grup, w sposób daleki od wymuszania i asymilacji”, por. L. Korpowicz, *Wielokulturowość i międzykulturowość w: U progu wielokulturowości. Nowe oblicza społeczeństwa polskiego*, Kempy, M., Kapciak, A., Łodziński, S., (red.), Instytut Socjologii Uniwersytetu Warszawskiego, Warszawa 1997, s. 64–65. Sądzę, że również w przypadku stylu myślowego Flecka jako odpowiednika danej kultury pojęcie międzykulturowości będzie bardziej odpowiednie.

²⁶ Na pozorną łatwość w definiowaniu wielokulturowości zwracają uwagę także inni autorzy, jak chociażby Julita Makaro i Kamila Dolińska. Badaczki wyróżniają cztery poziomy analizy problemu. Pierwszy poziom nawiązuje do definicji i dotyczy realnego zróżnicowania (różnorodność jest więc zjawiskiem demograficznym). Poziom drugi ma charakter świadomościowy i oznacza doświadczanie obecności innych kultur w codziennym życiu. Poziom trzeci zajmuje wielokulturowość polityczna lub multikulturalizm. Poziom czwarty, marketingowy, związany jest z kwestią wykorzystania wielokulturowości, zob. J. Makaro, K. Dolińska, *Wielokulturowość Europy Środkowo-Wschodniej. Kilka metodologicznych uwag o definiowaniu i badaniu zjawiska*, Sprawy Narodowościowe, 40, 2012, s. 87–100.

²⁷ M. Kubiszyn, *Edukacja wielokulturowa w środowisku lokalnym. Studium teoretyczno-empiryczne na przykładzie ośrodka „Brama Grodzka – teatr NN” w Lublinie*, Adam Marszałek, Toruń 2007, s. 19–22.

aktualność wspomnianej koncepcji i jej kompatybilność ze współczesną postacią zjawiska wielokulturowości.

Niewątpliwie warty podkreślenia jest wielokulturowy, a z punktu widzenia niniejszych rozważań, międzykulturowy charakter Lwowa, w którym dorastał i kształcił się Ludwik Fleck. Patrząc na historię Polski, można dostrzec, że wielokulturowość to stan bardzo częsty, ze szczególnym wyróżnieniem Polski Jagiellońskiej czy okresu międzywojennego.²⁸ Tyle że o tożsamości wspólnoty niekoniecznie decydowały względy etniczne czy kulturowe (np. wspólny język). Mieszczanie toruńscy w średniowieczu mówili po niemiecku, chociaż etnicznie nie można ich za Niemców. Zdecydowanie większą rolę odgrywały zatem względy polityczne. Różnorodnie etnicznie było Cesarstwo Austro-Węgierskie, gdzie wśród grup wyróżnić można Polaków, Czechów, Żydów; Węgrzy zatem mieli prawo nie czuć się gospodarzami we własnym kraju. Andrzej Szahaj pisze nawet, że „stan wielokulturowości był pewnym standardem w życiu większości krajów i państw na przestrzeni dziejów, zaś stan homogeniczności kulturowej lub etnicznej był czymś wyjątkowym”.²⁹ Ta swego rodzaju naturalność wielokulturowości zaczęła się rozmywać w XIX wieku pod wpływem tendencji nacjonalistycznych. Proces miał swoje uzasadnienie, powstawały wszak nowoczesne państwa narodowe dążące do unifikacji kulturowej. Na potrzeby wykazania związków stylu myślowego i rozumienia zjawiska wielokulturowości przyjęto wyjaśnienie tego terminu, jako współwystępowanie wielu kultur na określonym terytorium, z uwzględnieniem kryterium czasowego.

5. KONCEPCJA STYLU MYŚLOWEGO A FENOMEN WIELOKULTUROWOŚCI. KOMPATYBILNOŚĆ CZY WYKLUCZANIE SIĘ?

Fleck pisze w *Powstawaniu i rozwoju faktu naukowego*, że jeśli porówna się dwa style myślowe ze sobą, od razu będzie można dostrzec mniejsze lub większe różnice między nimi. Dla tego filozofa różnica między stylem myślowym fizyków i biologów nie będzie znacząca; dużo większa będzie między stylem myślowym fizyka i filologa. W przypadku stylów europejskiego i chińskiego lekarza różnica wydaje się być nie do zniwelowania. W zasadzie można mówić o odcieniach stylu, o jego odmianach oraz o różnych stylach. Im większą różnicę między dwoma stylami można dostrzec, tym mniejszą widać szansę na wymianę myśli. Jeśli zaistnieją międzykolektywne stosunki, to znaczy, że będą wykazywać wspólne rysy, bez względu na różnice, na oso-

²⁸ Teza, bardzo często słyszana zresztą, że współczesna Polska jest krajem jednolitym kulturowo, jest dość osobliwa. Odmienne kultury możemy obserwować na scenie politycznej, jak sądzę; odmienność jest zauważalna po przeciwnych jej stronach (np. liberalizm a konserwatyzm).

²⁹ A. Szahaj, *Zmierzch multi-kulti*, dostępny w internecie: rp.pl [dostęp: 11.04.2023r.].

bliwości odnośnych kolektywów. Zasady obcego kolektywu odczuwa się, jeśli w ogóle się je spostrzega jako arbitralne, a ewentualne ich uzasadnienie jako *petitio principii*. Obcy styl myślowy wydaje się być mistyczny, odrzucone przez niego pytania są często uważane za najważniejsze, wyjaśnienia za nieprzekonujące albo chybione. Jeśli przyjmiemy, że każdy styl myślowy jest osobną, inną kulturą, porozumienie między stylami myślowymi czy kolektywami jako ich nośnikami będzie trudne, jeśli nie niemożliwe. Skoro myśl ulega przekształceniu tak mocno, że nie jest zrozumiała dla odbiorcy z innego kolektywu, to czy oznacza to, że koncepcja stylu myślowego wyklucza wielokulturowość? Jeśli przyjąć za wiążącą tę definicję wielokulturowości, która uznaje współistnienie różnych grup kulturowych na obszarze terytorium jednego państwa, funkcjonowanie takiej wielości kultur, przynajmniej z jednej strony, wydaje się być problematyczne w świetle tej koncepcji. Koncepcja Flecka wymaga modyfikacji, aby dostosować ją do doświadczenia współczesnego człowieka, gdyż żyjący współcześnie człowiek doświadcza tego, że ludzie „przechodzą” między stylami myślowymi.

Jak powiada Fleck, styl myślowy pociąga za sobą gotowość do określonego widzenia. Skoro tak jest, zmiana kolektywu powoduje zmianę widzenia; zatem zmieniająca kolektyw jednostka widzi co innego. Tak więc, na przykład, pracownik naukowy może postrzegać swoje koleżanki jako partnerki naukowe. Jednak zmiana środowiska może spowodować zmianę optyki: wówczas kobieta może stać się wyłącznie obiektem gospodarczo-seksualnym.³⁰ Trudno w takiej sytuacji posądzać kogoś o hipokryzję. Jednostka doświadczająca takiej zmiany wierzy w pozornie sprzeczne tezy. Jednocześnie ważne wydaje się pytanie: czy tezy te są sprzeczne, czy są nierównoważne? Jako inny przykład można wskazać przypadek naukowca, który przez kilka roboczych dni tygodnia jest zdeklarowanym ewolucjonistą, w niedzielę zaś jego naukowe poglądy odsuwane są w cień na rzecz odmiennych koncepcji, którym on zdaje się wówczas oddawać.

Pojęcie stylu myślowego (paradygmatu) ma szerokie zastosowanie w obszarze analizy nauki czy historii nauki, natomiast nie ma takiego zastosowania na innych płaszczyznach, chociaż takie mu się czasem nadaje. Dużo trudniej jest go przenieść na płaszczyznę analizy zachowań kulturowych. W sytuacji, gdy dochodzi do kontaktów różnych społeczności, wytworzeniu różnych stylów myślowych może towarzyszyć jakaś potrzeba, realizacja jakichś wspólnych interesów. Ta potrzeba dotyczy takiego funkcjonowania, które może się dokonywać przy niezmiennym stylu myślowym, albo, jeśli to niemożliwe, przy wytwarzaniu pewnego rodzaju nadrzędnego stylu myślowego, który będzie ich jednoczył. Można zaryzykować tezę, że w obu stylach

³⁰ Z drugiej strony warto rozważyć, czy na pewno będziemy tu mówić o zmianie widzenia, czy może o eksponowaniu odmiennych funkcji. Nietrudno przecież wyobrazić sobie sytuację, w której partner z zespołu badawczego jest jednocześnie obiektem seksualnym (np. pracujące razem małżeństwa). Tak więc może bardziej będzie tu zdumienie, może dezaprobata, że tak też można patrzeć.

myślowych były jakieś wspólne przesłanki, stąd wytworzenie się jakiegoś nadrzędnego, jednoczącego stylu myślowego było w ogóle możliwe. Można także rozumieć to w inny sposób. A mianowicie: o wszystkim zadecydowały potrzeby tak silne, że utworzono nadrzędny styl myślowy mimo istotnych różnic, które to różnice spisano w tzw. „protokole rozbieżności”, jako swoim dokumentem zawierającym te kwestie, co do których nie osiągnięto porozumienia. Ta swego rodzaju dialektyka może być dwustronnie motywowana. Albo potrzebami o różnym charakterze, jak choćby praktycznymi, lub też może pochodzić z jakichś wspólnych przesłanek: wszyscy jesteśmy ludźmi, mamy jakieś wspólne założenia, różnimy się tylko obyczajami a obyczaje te nie są czymś, czego nie udało się przekroczyć/pokonać.

Być może wielokulturowość niemożliwa jest jako konstatacja teoretyczna. Wspólnistnieje wiele kultur, gdyż są one zdeterminowane przez style myślowe, ale niemożliwe jest funkcjonowanie jednostki w warunkach transkulturowych. Wielokulturowość może być wartością; cenimy różnice a jest to możliwe, gdy stajemy się w pewien sposób nadrzędni, aby patrzeć na świat społeczny i kulturowy jako na świat zróżnicowany i to zróżnicowanie traktować jako coś cennego. Nasuwa się zatem, chyba uzasadnione, pytanie o to, kto się myli: Ludwik Fleck czy teoretycy wielokulturowości? Faktem jest uczestnictwo w wielu kulturach, w wielu stylach myślowych, zatem izolacja myślowa i ścisła determinacja stylu myślowego nie ma miejsca i w tym rozumieniu Fleck raczej przecenił rolę stylu myślowego.

Nie jest także bez racji teza, że niektóre style myślowe mogą być przydatne do opisu fenomenu wielokulturowości, inne zaś nie. Nawiązując do metafory tygła (*melting pot*) można zaryzykować stwierdzenie, że pewne kultury mogą łączyć się ze sobą i tym samym współżyć z innymi na obszarze jednego państwa, inne tego nie potrafią, dlatego przypominają bardziej *salad bowl* (kultury żyją oddzielnie obok siebie, nie myśląc o asymilacji). Wracając jednak do metafory obrazującej kulturę łączącą się ze sobą, postawić można problem, czy nie dlatego mogą one się ze sobą mieszać, że są częścią czegoś nadrzędnego – jakiegoś ogólniejszego stylu myślowego?

Jak się mają wybitne jednostki do zmian w obrębie kultury i wielokulturowości? Wydaje się, że można je traktować jako pasterzy (wehikuły) wielokulturowości, tyle tylko, że i to byłoby niezgodne z tezą Flecka o ściśle izolującej funkcji stylu myślowego. Jednostka wybitnie twórcza jest u niego kwestią marginalną.³¹ W historii kultury, tak naprawdę, rzadko jakkolwiek transformacja kulturowa była spowodowana przez jednostkę. Kuhn chociażby przypomina, że rewolucja kopernikańska nie była wynikiem działań sa-

³¹ W zasadzie wspomina jedynie, że to zagadnienie odrębne; zob. L. Fleck, *Powstanie i rozwój faktu naukowego*, s. 68. Dokonuje przy tym indywidualizacji jednostki, reprezentantki pewnych funkcji społecznych, osłabiając ją poprzez wahanie terminologiczne: „wyjątkowo twórcze jednostki czy może raczej twórcze momenty”; zob. M. Cyzman, *Kolektyw czy jednostka? Kto tworzy teorie (nie tylko) naukowe? O pewnym szczególe w koncepcji Ludwika Flecka*, *Er(r)go. Teoria-Literatura-Kultura*, 45, 2022, s. 110.

meo Kopernika a raczej tego, że środowisko naukowe zareagowało na rezultaty pracy polskiego astronoma i odkrywcy. Właśnie to obrazują słowa Flecka: „działa kolektyw myślowy, nie jednostka” co z kolei stanowi argument potwierdzający tezę, że myśli Ludwika Flecka są ze sobą sprzeczne.

Wielokrotnie także u autorów odwołujących się do myśli Flecka i komentujących jego twierdzenia, można odnaleźć tezy, że powstanie stylu myślowego poprzedzają uwarunkowania kulturowe. W klasycznej koncepcji prawdy mówimy o relacji teorii ze światem (z rzeczywistością). Styl myślowy należy umieścić gdzieś pomiędzy. Gdyby kultura była istotnym elementem tejże relacji, ona także musiałaby zostać gdzieś na tej linii umieszczona. A jeśli kultura, to niewykluczone że i inne ogniwa pośrednie.

Istotne jest pytanie o możliwość bycia człowiekiem „spoza stylu myślowego”. Czy wielokulturowość to styl myślowy – pewnego rodzaju metastyl? Ku temu wiodły zapatrywania w trakcie niniejszych rozważań. Zarazem jednak interpretacja stwierdzenia: „można być uczestnikiem wielu stylów myślowych” prowadziła do upewniania się, że chyba jednak istnieje możliwość bycia w jakiś sposób poza stylem myślowym. W związku z tym wydaje się, że style myślowe nie pełnią aż tak determinującej roli, jak przedstawiał to Ludwik Fleck. Albo więc możliwe jest funkcjonowanie poza stylem, albo jest to niemożliwe – wówczas jest się przedstawicielem innego stylu (nadrzędnego) myślowego (metastylu). Bardziej należy skłonić się ku opcji, że jest się przedstawicielem metastylu i być może zadać pytanie dodatkowe; czy przedstawiciele dowolnych kultur mogą wchodzić w interakcję, innymi słowy, czy mogą one taki metastyl tworzyć? Niewykluczone bowiem, że są kultury nie nadające się do koincydencji i niewykluczone również, że dziś właśnie mamy do czynienia z taką sytuacją – należy mieć tu na myśli relację: kultura wschodnia-kultura zachodnia.

Do koncepcji stylu myślowego podchodzić można z różnych punktów widzenia i szukać punktów wspólnych, bądź rozbieżnych, z ideą wielokulturowości. Uważam jednak, nawiązując do pytania problemowego ze wstępu do niniejszego artykułu, że nie można zawyrokować jednoznacznie: obie koncepcje są kompatybilne ze sobą/obie się całkowicie wykluczają. Nie można tego uczynić, gdyż wydaje się, że koncepcja stylu myślowego Ludwika Flecka jest nieadekwatna empirycznie, w wielu punktach jest ze sobą sprzeczna, przez co może tracić na aktualności i o ile w kwestiach naukowych (style naukowe) mogłaby się bronić, tak na innych płaszczyznach wykazuje wiele trudności. Zrozumiała będzie zatem obawa, że koncepcja Flecka nie jest ona przydatna do opisu fenomenu wielokulturowości.

Takie konkluzje mogą wydawać się z drugiej strony zbyt mocne, gdyż Fleck nie kwestionuje tego, że jednostka może uczestniczyć w wielu niegroźnych [kompatybilnych] stylach. Wyklucza jedynie to, że jednostka może być członkiem kolektywów o wrogich [niewspółmiernych] stylach np. jednocześnie ewolucjonista i kreacjonista. Zatem kultury z gruntu wrogie nie mogą

koegzystować pokojowo, a ich przedstawiciele, tj. jednostki, wchodzić w relacje, chyba, że realizują wspólnie podzielaną tak wysoką wartość, że odkładają na bok to, co je dzieli.³²

Rozważania, które zostały podjęte w tym punkcie artykułu, służyły nadaniu rysu krytycznego Fleckowskiej koncepcji. Z pewnością w obliczu przedstawionych argumentów nasuwa się wniosek i odpowiedź na wątpliwość postawioną w tytule – mianowicie, czy styl myślowy jest kompatybilny z fenomenem wielokulturowości, czy się z nim wyklucza. Wskazywane tutaj niekonsekwencje przemawiać mogą za tezą, że mamy do czynienia z wykluczeniem się obu koncepcji. Ujawnione sprzeczności nie przesądzają jednak o tym, że styl myślowy nie nadaje się do opisu i zrozumienia fenomenu wielokulturowości. Przeciwnie, może być doskonałym narzędziem do takich zabiegów. Ta właśnie teza będzie ostatecznie argumentowana w zakończeniu.

6. PODSUMOWANIE

W podsumowaniu artykułu konieczne staje się nawiązanie do problemów zasygnalizowanych we wstępie oraz do sformułowanej tam tezy. Argumentacja na rzecz tezy, że koncepcja stylu myślowego jest adekwatnym narzędziem do opisu oraz do rozumienia wielokulturowości, powinna, w mojej ocenie, skupiać się wokół trzech głównych zagadnień/kluczy: różnicy, nieprzenikalności/niewspółmierności oraz czasowości.

Zróznicowanie jest niezbędnym warunkiem powstawania i funkcjonowania stylów myślowych. Ludwik Fleck często zwraca uwagę na ich wielość i różnorodność, także w aspekcie diachronicznym. Istotne jest zatem i tutaj dostrzeżenie różnicy.³³ Socjologia myślenia może owocnie badać tak ważne zjawiska, jak propaganda, działanie autorytetu, rola naśladownictwa, współpraca i konkurencja intelektualna, drogi i sposoby rozprzestrzeniania się poglądów. Bada sprawę wprowadzania w pewien styl myślenia, sprawę odgraniczania się kolektywów i ich wewnętrznego organizowania. Należy wziąć tu również pod uwagę, że w kręgu zainteresowania socjologii myślenia umieszcza Fleck sprawę charakterystycznej struktury różnych kolektywów: kolektywu myśli naukowej, kolektywu życia codziennego, psychologii pewnych zawodów, pewnych klas, czy innych grup. Ważne słowa padają na końcu artykułu *Patrzeć, widzieć, wiedzieć*. Fleck pisze tam, że zamiast światopoglądu, który ulega ustawicznym zmianom i jest zależny od chwili i miej-

³²Przywołać można przykład dość zabawny, lecz i, wydaje się, adekwatny. Oto kibice zwaśnionych drużyn wspólnie są gotowi ratować potrąconą przez motocyklistę starszą kobietę, ale poza tym incydentem będą pałać do siebie szczerą nienawiścią.

³³L. Fleck, *Patrzeć, widzieć, wiedzieć*, w: *Style myślowe i fakty. Artykuły i świadectwa*, op. cit., s.184.

sca, socjologia myślenia da pogląd na mechanizm powstawania światopoglądów. W związku z tym „zamiast tego, co ludzi dzieli, wskaże im to, co będąc wszystkim wspólne zbliża ich do siebie”.³⁴ Nasuwa się zatem pytanie, czy znając mechanizmy powstania odmiennych sposobów myślenia, odmiennych kultur, nie można uzyskać płaszczyzny porozumienia. W moim odczuciu – można, chociaż przy niewspółmierności to zadanie będzie trudne.

Biorąc pod uwagę powyższe zagadnienia, podkreślić należy ścisły związek stylu myślowego i fenomenu wielokulturowości. Jeśli różnica jest podstawą ich wykształcenia się i funkcjonowania, to koncepcja stylu myślowego pozwoli na zrozumienie wielokulturowości.

Obydwa określenia, styl myślowy i wielokulturowość, zdają się wiązać ze sobą, i to dość ściśle. Myślę, że warto uwzględnić ich swego rodzaju stopniowalność. Nieprzenikalność natomiast, tak charakterystyczną w odniesieniu do stylu myślowego, zastosować do braku porozumienia w drobnej (wielokulturowej) kwestii i odróżnić od niewspółmierności, gdzie porozumienie nigdy nie będzie możliwe. Z pewnością nie należy traktować ich synonimicznie.

Nie mniej istotne jest wydobycie na światło dzienne niewspółmiernego aspektu stylów myślowych. Koncepcja stylu myślowego nie wyklucza kontaktów międzykolektywnych, zdaje się zakładać jedynie, że w przypadku stylów niewspółmiernych kontakty te, jak i osiągnięcie porozumienia, mogą być utrudnione. Jednak koncepcja stylu myślowego może pomóc w wysunięciu konkluzji, że nie w każdych okolicznościach osiągnięcie takiego porozumienia jest możliwe i należy potraktować to jako fakt. W tym postrzegalabym aktualność i znaczenie koncepcji Flecka, który, osadzając ją w socjologii myślenia, podkreślał znaczenie tej dziedziny dla wytlumaczenia zagadnienia porozumienia i braku porozumienia się ludzi. To, że Fleck budując swoją koncepcję przestrzegał przed stosami, które zapalą się zawsze, gdy członkowie kolektywu będą zwalczać inne style, świadczy o wyjątkowej aktualności jego myśli. Zderzenie odmiennych kultur, które odmienności nie chcą potraktować jako szansy, prowadzi do poważnych konsekwencji.

Okazuje się, że ją także należy rozpatrywać w aspekcie czasowych. Fleck mówi o tym bardzo wyraźnie, przytaczając historię pojęcia kiły. Obserwującemu losy idei nasuną się odrębne zjawiska z historii poznania. Wiele teorii np. przeżywa dwie epoki: najpierw klasyczną, w której „wszystko się zgadza”, a następnie drugą, w której dopiero pokazują się wyjątki. Albo okazuje się, że niektóre idee pojawiają się o wiele wcześniej niż ich racjonalne podstawy, są więc zupełnie od tych ostatnich niezależne. Właściwy opis dziejów jakiegś dziedziny jest rzeczą trudną, przestrzega Fleck; gdyż składa się na nią wiele krzyżujących się i przenikających się wzajemnie myślowych ciągów rozwojowych. Niektóre pojęcia znikają, inne powstają. Odbyma się to bez względu

³⁴ Ibidem.

na rzeczowe lub logiczne argumenty: zarzucono np. zagadnienie kamienia filozoficznego, nie dlatego, że poszukiwania były bez wyniku, ale dlatego, że zagadnienie nie zgadzało się ze stylem myślowym kolejnych czasów. W mojej ocenie kompatybilność stylu myślowego z fenomenem wielokulturowości jest tutaj bardzo widoczna.

W nawiązaniu do tezy głównej w artykule, która głosiła, że koncepcja stylu myślowego jest użytecznym i adekwatnym narzędziem do opisu wielokulturowości, należy wyekspozować jej zasadność: koncepcja ta jest użytecznym narzędziem do opisu wielokulturowości i z pewnością nie można jej odmówić braku aktualności. W artykule wykazywałam nasuwające się wątpliwości i nieścisłości kryjące się w konstrukcji i działaniu stylu myślowego a które wydały się istotne. Podawałam w tym celu wiele przykładów z życia codziennego. Nieścisłości te jednak nie mogły zaważyć na ogólnej ocenie użyteczności koncepcji, gdyż nie dotyczyły kwestii zasadniczych, które przed chwilą zostały zasygnalizowane. Tym samym znaczenie Flecka jako filozofa nauki, jako prekursora społecznej natury poznania i jako autora „stylu myślowego” powinno rosnać.

BIBLIOGRAFIA

- T. Biblikiewicz, *Uwagi nad artykułem Ludwika Flecka »Nauka a środowisko«*, w: L. Fleck, *Powstanie i rozwój faktu naukowego. Wprowadzenie do nauki o stylu myślowym i kolektywie myślowym*, przeł. M. Tuskiewicz, Wydawnictwo Lubelskie, Lublin 1986.
- M. Ciesielska, P. Jarnicki, *Ludwik Fleck – mikrobiolog i filozof*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2021.
- M. Czyżman, *Kolektyw czy jednostka? Kto tworzy teorie (nie tylko) naukowe? O pewnym szczególnie w koncepcji Ludwika Flecka*, *Er(r)go. Teoria-Literatura-Kultura*, 45, 2022.
- B. Dziobkowski, *Realizm wewnętrzny a relatywizm*, *Przegląd Filozoficzny – Nowa Seria*, 3, 2006.
- L. Fleck, *Odpowiedź na uwagi Tadeusza Bilikiewicza*, w: tenże, *Psychosocjologia poznania naukowego. Powstanie i rozwój faktu naukowego oraz inne pisma z filozofii prawa*, Z. Cackowski, S. Symotiuk (red.), Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, Lublin 2006.
- L. Fleck, *Patrzeć, widzieć, wiedzieć*, w: S. Werner, C. Zittl, F. Schmaltz (red.), *Style myślowe i fakty. Artykuły i świadectwa*, Wydawnictwo Instytutu Filozofii i Socjologii Polskiej Akademii Nauk, Warszawa 2007.
- L. Fleck, *Powstanie i rozwój faktu naukowego*, w: tenże, *Psychosocjologia poznania naukowego. Powstanie i rozwój faktu naukowego oraz inne pisma z filozofii prawa*, Z. Cackowski, S. Symotiuk (red.), Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, Lublin 2006.
- L. Fleck, *Powstanie i rozwój faktu naukowego. Wprowadzenie do nauki o stylu myślowym i kolektywie myślowym*, przeł. M. Tuskiewicz, Wydawnictwo Lubelskie, Lublin 1986.
- R. W. Gryglewski, *Ludwik Fleck*, w: *Giganci nauki. Biogramy*, dostępny w internecie: <https://gigancinauki.pl/gn/biogramy/82990,Fleck-Ludwik.html#> [dostęp: 20 listopada 2024 r.].
- P. Jarnicki, *Antytotalitarne motywy twórczości Ludwika Flecka*, w: B. Płonka-Syroka, A. Kaźmierczak (red.), *Nauka w kontekście wzorców kultury*, Wydawnictwo DiG, Warszawa 2011.
- P. Jarnicki, *Ludwika Flecka nauka bez prawdy?*, *Przegląd Filozoficzny – Nowa Seria*, 2, 2010.

- L. Korpowicz, *Wielokulturowość i międzykulturowość*, w: M. Kempy, A. Kapciak, S. Łodziński (red.), *U progu wielokulturowości. Nowe oblicza społeczeństwa polskiego*, Oficyna Naukowa, Warszawa 1997.
- M. Kubiszyn, *Edukacja wielokulturowa w środowisku lokalnym. Studium teoretyczno-empiryczne na przykładzie ośrodka „Brama Grodzka – teatr NN” w Lublinie*, Wydawnictwo Adam Marszałek, Toruń 2007.
- W. Kymlicka, *Współczesna filozofia polityczna*, przeł. A. Pawelec, Aletheia, Warszawa 2009.
- J. Makaro, K. Dolińska, *Wielokulturowość Europy Środkowo-Wschodniej. Kilka metodologicznych uwag o definiowaniu i badaniu zjawiska*, *Sprawy Narodowościowe*, 40, 2012.
- H. Mamzer, *Tożsamość w podróży. Wielokulturowość a kształtowanie tożsamości jednostki*, UAM, Poznań 2002.
- M. Mead, *Kultura i tożsamość. Studium dystansu międzykulturowego*, przeł. J. Hołówka, PWN, Warszawa 2000.
- T. Möller, *Uwagi krytyczne na temat pojęć kolektyw myślowy, styl myślowy i wymiana myśli*, w: S. Werner, C. Zittl, F. Schmaltz (red.), *Style myślowe i fakty. Artykuły i świadectwa*, Instytut Filozofii i Socjologii Polskiej Akademii Nauk, Warszawa 2007.
- H. Putnam, *Wiele twarzy realizmu i inne eseje*, przeł. A. Grobler, PWN, Warszawa 1998.
- M. Rydlewski, *W stronę teorii i historii kultury. Spojrzenie na teorię stylów myślowych Ludwika Flecka*, *Etnografia Polska*, 1–2, 2009.
- W. Sady, *Fleck. O społecznej naturze poznania*, Prószyński i S-ka, Warszawa 2000.
- R. Sojak, *Czy spór o relatywizm można rozwiązać empirycznie?*, w: P. Bytniewski, M. Chałubiński (red.), *Teoretyczne podstawy socjologii wiedzy*, t. 1, Wydawnictwo UMCS, Lublin 2006.
- A. Szahaj, *Zmierzch multi-kulti*, dostępny w internecie: rp.pl [dostęp: 11.04.2023 r.].

LUDWIK FLECK'S CONCEPT OF THOUGHT STYLE AND THE UNDERSTANDING OF MULTICULTURALISM

ABSTRACT

The concept of thinking style finds its application in the scientific world. The author himself gives such an application of it, drawing exoteric and also esoteric circles within it and situating specialists in a particular field and laymen there. However, it seems that the concept of thinking style can also be successfully applied in the field of culture, especially with regard to the coexistence of many cultures side by side and their interactions. The purpose of undertaking the considerations in this article, in addition to introducing the figure of a resident of Galicia of the early twentieth century itself, is to highlight the relationship of thought style and the phenomenon of multiculturalism. The juxtaposition of thought style and multiculturalism, a perhaps difficult and novel task, seems important, as it may demonstrate the enduring relevance in today's world of Fleck's concepts of almost a century ago.

Keywords: thought style, multiculturalism, thought collective, sociology of knowledge.

O AUTORCE — doktorantka w Szkole Doktorskiej Uniwersytetu Rzeszowskiego. Obszary badawcze to filozofia nauki oraz filozofia religii. Najważniejsze podejmowane problemy badawcze: relacja nauki i filozofii, nauki i religii oraz rola koncepcji stylu myślowego Ludwika Flecka w organizacji i funkcjonowaniu społeczeństw wielokulturowych. Absolwentka studiów magisterskich z zakresu komunikacji międzykulturowej oraz współorganizatorka konferencji naukowych o dialogu międzykulturowym. Zagadnienia przedstawiane w dysertacji dotyczą religijnego i eschatologicznego pluralizmu w ponowoczesności.

E-mail: anetar@dokt.ur.edu.pl

Jacek Gurczyński

MANIPULACJA EMOCJAMI I DANE AFEKTYWNE W EPOCE CYFROWEJ

<https://doi.org/10.37240/FiN.2025.13.05>

STRESZCZENIE

Tekst podejmuje krytyczną analizę manipulacji emocjonalnej w mediach społecznościowych, ujmując ją jako element szerszej infrastruktury cyfrowego kapitalizmu afektywnego. Punktem wyjścia są empiryczne przykłady ingerencji w emocje użytkowników – eksperyment Facebooka z 2014 roku oraz działalność Cambridge Analytica – które ujawniają zdolność platform cyfrowych do modulowania nastrojów, decyzji i zachowań politycznych przy użyciu algorytmów, danych osobowych i psychograficznego mikrotargetowania. Tekst pokazuje jednak, że narracja o „wszechmocy” tych technologii bywa przesadzona, a modele osobowości i predykcji emocjonalnych są epistemologicznie kruche, uproszczone i silnie kontekstualne.

W dalszej części analiza przesuwana się z poziomu spektakularnych skandali ku codziennym praktykom afektywnym, takim jak użycie emoji, emotikonów i reakcji platformowych. Zostają one zinterpretowane jako narzędzia standaryzacji emocji, które jednocześnie umożliwiają ekspresję afektywną i redukują jej złożoność, czyniąc ją podatną na agregację, analizę sentymentu i komodyfikację. Odwołując się do filozofii procesualnej Whiteheada, teorii performatywności Butler, koncepcji pracy emocjonalnej Hochschild oraz krytyki kapitalizmu nadzoru Zuboff, tekst argumentuje, że emocje w mediach społecznościowych nie są jedynie wyrażane, lecz współprodukowane przez architekturę platform.

Emoji i systemy analizy sentymentu zostają ukazane jako elementy biopolityki emocji, w której afekt staje się surowcem ekonomicznym, a użytkownicy – afektywnymi pracownikami platform. Jednocześnie tekst wskazuje na granice algorytmicznego ujęcia emocji: ich kontekstualność, wieloznaczność i relacyjny charakter wymykają się pełnej automatyzacji. W zakończeniu zaproponowana zostaje alternatywna perspektywa badawcza, inspirowana koncepcją transindywidualności i komunitaryzmu emocjonalnego, która traktuje afekt nie jako towar, lecz jako zasób relacyjny, polityczny i etyczny, zdolny do budowania solidarności i oporu wobec instrumentalnej logiki cyfrowego kapitalizmu.

Słowa kluczowe: manipulacja emocjonalna; media społecznościowe; emoji; analiza sentymentu; afekt; praca emocjonalna; kapitalizm nadzoru; dane osobowe;

algorytmy; psychograficzne mikrotargetowanie; transindywidualność; biopolityka emocji.

MANIPULACJA EMOCJONALNA W MEDIACH SPOŁECZNOŚCIOWYCH

Przykładem manipulacji emocjonalnej jest eksperyment przeprowadzony na Facebooku na grupie 689 000 użytkowników przez naukowców z Cornell University oraz University of San Francisco, w którym manipulowano treściami w newsfeedach, by kontrolować emocje, na jakie byli wystawieni użytkownicy (Kramer et al., 2014). Eksperyment ten budził poważne zastrzeżenia etyczne, m.in. ze względu na brak zgody uczestników, ale stał się dowodem na zdolność mediów społecznościowych do manipulowania myślami, uczuciami i emocjami. Współcześnie tę moc manipulacji wykorzystwała firma Cambridge Analytica, która bez zgody użytkowników i Facebooka zebrała dane osobowe 50 milionów użytkowników, by manipulować ich zachowaniami wyborczymi. Działania te zostały powiązane z kampanią prezydencką Donalda Trumpa w USA oraz z kampanią Brexit w Wielkiej Brytanii. Cambridge Analytica opracowała oprogramowanie zdolne do przewidywania i emocjonalnego wpływania na decyzje wyborcze, znane jako psychograficzne mikrotargetowanie. Jak przyznał sygnalista Christopher Wylie: „Wykorzystaliśmy Facebooka do zebrania profili milionów ludzi. Stworzyliśmy modele, które wykorzystywały to, co wiedzieliśmy o nich, by celować w ich wewnętrzne demony. Na tym opierała się cała firma.” (Graham-Harrison & Cadwalladr, 2018).

Dzięki algorytmom wykrywającym cechy osobowości wyborców, które bazowały na modelu pięciu wielkich czynników (OCEAN: otwartość na doświadczenia, sumiennność, ekstrawersja, ugodowość, neurotyczność), tworzono przekazy mające na celu nakłonienie do głosowania na Trumpa lub za Brexitem. Przykładowo, osoby o wysokim poziomie neurotyczności, uważane za podatne emocjonalnie, były szczególnie intensywnie targetowane. Kampania wykorzystywała tzw. „Dark Posty” — spersonalizowane reklamy widoczne tylko dla wybranych użytkowników. Krytycy teorii cech osobowości, jak Mischel (1968) czy Kelly (1970), zwracają uwagę, że testy osobowości są subiektywne, niskiej wiarygodności i oparte na samowiedzy uczestników, a wiele cech jest kontekstualnych i zmiennych. Krytycy społeczni postulują, że osobowość jest konstruktem społecznym i kategorią reifikującą (Ellis, Tucker, 2021). Ponadto, choć Cambridge Analytica chwaliła się przewidywaniem zachowań wyborczych, badania wykazały, że profile wyborców słabo korelują z ich poglądami politycznymi (Chen & Potenza, 2018). Nawet jej były pracownik Rick Tyler przyznał, że psychograficzne modele były niewiarygodne. Niektórzy analitycy uważają, że Cambridge Analytica wyolbrzymiła

swoj wpływ na wybory. Taggart (2017) wskazał, że podejście psychologiczne firmy nie było realnie stosowane przez kampanię Trumpa, a opowieść o przełomie w technologii politycznej jest raczej elementem walki o wpływy między konsultantami.

Pytanie o to, czym jest informacja osobista i prywatna, staje się kluczowe w epoce cyfrowej. Dane powinny być traktowane jako informacje o sobie, które są cyfryzowane i przekształcane w towary na rynku cyfrowym? Ellis (2018) sugeruje, że użytkownicy mediów społecznościowych coraz częściej rozumieją dane jako bezosobowe, a nie osobiste. Problem dotyczy własności tych danych, które są wytwarzane w ramach tego, co Zuboff (2019) nazywa „nadwyżką behawioralną” w kapitalizmie nadzoru. Proces udostępniania informacji osobistych jest wynikiem narzuconej przez media społecznościowe umowy społeczno-technologicznej, którą użytkownicy często akceptują nieświadomie, zgadzając się na regulaminy i warunki. Jednak czy rozumieją, co oznacza ta wymiana? Czy następuje rozczłonkowanie jaźni na części sprzedawalne i fragmentowane? Ellis (2018, s. 21) opisuje, jak proces „utrąty dynamizmu” danych osobistych prowadzi do ich „zamrożenia” i kategoryzacji w celach cyfrowego kapitalizmu. Paradoksalnie, choć statusy na WhatsAppie czy posty z emotikonami mogą wydawać się bardziej osobiste i dynamiczne niż twarde dane osobowe, to jednak wszystkie te informacje są zbierane i wykorzystywane w ekosystemach cyfrowych. Manipulacja emocjami w mediach społecznościowych ma charakter systemowy i wielowymiarowy. Eksperymenty takie jak ten przeprowadzony przez Kramera i współpracowników (2014) pokazują, jak łatwo jest wpływać na uczucia milionów użytkowników poprzez selektywną prezentację informacji. Tego typu praktyki nie są już tylko eksperymentami naukowymi, lecz stały się integralnym elementem działań korporacji i kampanii politycznych.

Cambridge Analytica, wykorzystując dane osobowe, psychograficzne profile i zaawansowane algorytmy, dążyła do perfekcyjnego targetowania wyborców – dotarcia do ich „wewnętrznych demonów”, czyli najbardziej wrażliwych emocjonalnych punktów. Jednak, jak zauważają krytycy, efektywność takich działań jest mocno przesadzona. Psychologiczne teorie cech osobowości są często zbyt uproszczone i kontekstowo zależne, co osłabia trafność przewidywań (Chen & Potenza, 2018; Taggart, 2017). Mimo to, potencjał manipulacji pozostaje ogromny, szczególnie w sytuacjach, gdy użytkownicy nie są świadomi zakresu i mechanizmów zbierania oraz wykorzystywania ich danych. Współczesne systemy cyfrowe tworzą ekosystemy danych, w których jednostki są nie tylko konsumentami, lecz przede wszystkim źródłem danych – surowcem poddawanym obróbce i wykorzystaniu do celów komercyjnych i politycznych (Zuboff, 2019; Nash, 2016).

Ellis (2018) podkreśla, że dane generowane przez użytkowników w mediach społecznościowych tracą swoją osobistą, dynamiczną naturę, stając się „zamrożonymi” reprezentacjami, które są kategoryzowane i wykorzystywane

w systemach kapitalizmu cyfrowego. W świetle filozofii procesualnej Whiteheada (2011) można to rozumieć jako transformację żywej, dynamicznej tożsamości w statyczne „aktualne byty” (actual entities), które oddzielają użytkownika od jego własnego „ja”. Taki proces prowadzi do alienacji, gdzie jednostka nie ma kontroli nad swoimi danymi, które stają się towarem. Anonimowość i bezosobowość danych sprzyjają utracie poczucia indywidualnej tożsamości i wzmacniają lęki związane z brakiem kontroli (Nash, 2016).

ZAKODOWANE EMOCJE: EMOJI, PRACA AFEKTYWNA I KOMODYFIKACJA PODMIOTOWOŚCI W KAPITALIZMIE CYFROWYM

Emotikony i emoji są próbą symbolicznego wyrażenia emocji w formie cyfrowej. Służą one uproszczeniu i standaryzacji komunikacji emocjonalnej w mediach społecznościowych, ale jednocześnie są narzędziem, które ułatwia agregację danych i ich komercyjne wykorzystanie (Ellis & Tucker, 2021). Nash (2016) analizuje to z perspektywy ontogenezy Simondona, gdzie indywidualizacja jest procesem ciągłym, a emoji są elementem umożliwiającym modulację i transmisję afektów w cyfrowych systemach. Jednak mechanizm ten jest wykorzystywany do utrzymania jednostki w stanie „stacycznej indywidualności”, co wzmacnia alienację i lęk. Pierwotną funkcją emotikonów – jak przypominają autorzy – było zarządzanie niejednoznacznością komunikacji tekstowej w warunkach braku komunikacji niewerbalnej (Fahlman, n.d.). Pojawienie się sekwencji :-) miało zabezpieczyć wypowiedzi przed błędną interpretacją, zwłaszcza w przypadku ironii czy żartu. Jednak ich funkcja ewoluowała – dzisiaj emotikony i emoji pełnią również funkcję zastępczą wobec emocji, a nawet stają się ich substytutem (Ellis & Tucker, 2020, s. 24).

Użytkownicy przyznają, że emotikony „pomagają” im w wyrażaniu emocji, których nie potrafią lub nie chcą ująć słowami. Są narzędziem redukcji złożoności i jednocześnie wsparcia emocjonalnego – wizualnym skrótem (Ellis, Tucker, 2020, s. 25). Jest to jednak wsparcie, które ogranicza – selekcjonuje i porządkuje afekty według dostępnych kategorii, często z wyłączeniem subtelności czy ambiwalencji. Ważnym aspektem analizy jest rozpoznanie, że emoji mogą być używane również do ukrywania, a nie tylko wyrażania emocji. Jeden z uczestników badania przyznaje, że korzysta z emotikonów i statusów, ponieważ w świecie online nie musi „brać tego na serio” (Ellis, 2018, s. 26). W odróżnieniu od interakcji twarzą w twarz, w których emocje są „pokazywane” ciałem, w świecie cyfrowym użytkownik może nimi zarządzać, je modulować i zasłaniać. Z perspektywy dramaturgicznej teorii emocji (Goffman, 1967; Hochschild, 1983), mamy tu do czynienia z cyfro-

wym odpowiednikiem „face-work” i „emotion work” – pracy nad emocjonalnym wizerunkiem. Emotikony stają się elementem fasady, strategią autoprezentacji, która umożliwia afektywną kontrolę i dystans wobec emocjonalnej autentyczności (Ellis & Tucker, 2015).

W latach 90. komunitarystyczne ideologie kreowały wizję mediów cyfrowych jako przestrzeni wspólnot, które miały oferować poczucie przynależności i wspólnotowego celu (Williams, 1961; Rheingold, 1994). Jednak współczesne doświadczenia pokazują, że rzeczywistość mediów społecznościowych jest bardziej ambiwalentna i często przyczynia się do izolacji i lęku (Morozov, 2013; Nash, 2016). Emocjonalna złożoność codziennych doświadczeń, takich jak spóźnienie się na autobus, może być w praktykach cyfrowych komunikacji zredukowana do prostego, wizualnego znaku – emoji. Jak zauważają Ellis i Tucker, zamiast wyrażać wielowarstwowe uczucia (frustrację, złość, bezsilność, wstyd), użytkownik mediów społecznościowych sięga po graficzny symbol, który reprezentuje stan afektywny w uproszczony sposób – (Ellis, 2018, s. 22). Zgodnie z procesualną filozofią Whiteheada, takie doświadczenie emocjonalne może być postrzegane jako „actual occasion” – konkretna aktualizacja cielesno-uczuciowego momentu. Jednak poprzez użycie emoji, dochodzi do „prehensji” (prehension) – selektywnego przyswojenia wybranych aspektów tego doświadczenia, a pominięcia innych (Whitehead, 1978). Emoji pełni funkcję „pozytywnej prehensji”, przekształcając dynamiczne doświadczenie w ustabilizowaną jednostkę znaczenia w systemie cyfrowym (Ellis, 2018, s. 23). Jednocześnie, wybór emoji nie jest świadomym, indywidualnym aktem decyzyjnym, lecz wynikiem strukturalnych ograniczeń infrastruktury platformy cyfrowej. Możliwości ekspresji są filtrowane przez algorytmy, interfejsy i normy komunikacyjne platform (Bucher, 2018). Użytkownik dokonuje wyboru spośród dostępnych reprezentacji, które jednocześnie upraszczają i standaryzują ekspresję afektywną.

Z perspektywy filozofii Judith Butler (2004), można odczytać emoji jako element performatywny – narzędzie powtarzalnych aktów, które tworzą, a nie tylko reprezentują emocje. Użytkownik, umieszczając emoji 😞, nie tyle „pokazuje”, co „tworzy” melancholijną tożsamość w ramach danej platformy. Jest to akt konstruowania siebie w ramach dostępnych znaków afektywnych. Emocja staje się tutaj efektem praktyki semiotycznej, wpisanej w ramy technologiczne. Jak pokazują badania neuropsychologiczne (Churches et al., 2014), widok emoji aktywuje te same obszary mózgu, które odpowiadają za reakcje emocjonalne na prawdziwe twarze. Jednakże, nie są to dokładnie te same procesy – np. brak aktywacji w rejonie odpowiedzialnym za rozpoznawanie twarzy (fusiform gyrus). To sugeruje, że mamy do czynienia z nowym typem kodu bio-kulturowego, który formuje nowe sposoby doświadczania afektów – oparte na estetyce i rytuałach cyfrowej komunikacji (Yuasa et al., 2006; Ellis, Tucker, 2020, s. 27). W efekcie, ekspresja afektywna w mediach społecznościowych nie tylko oddziałuje na innych, lecz również przekształca

podmiot – emocje są kodowane, selekcyjonowane i wyrażane w sposób dostosowany do reguł platformy. Jest to forma komodyfikacji emocji, a zarazem ich algorytmicznej dramaturgii.

Symbole wyłaniają się z wielości pragnień – były to pragnienia wyjaśniania, obiektywizowania, zwodzenia, ale również pragnienia utrwalania i kwalifikowania afektu, w ramach których prehensje zostają odarte ze swojej dynamicznej złożoności. Działanie to jest często wykorzystywane, zwłaszcza przez korporacje, w celu zwiększenia liczby kliknięć, polubień, znajomych, użytkowników i powiązanych z tym ekonomii. Stark i Crawford (2015) widzą emoji jako kanał pracy afektywnej. Sugerują, że znaki emoji ucieleśniają i reprezentują „napięcie między afektem jako ludzkim potencjałem a jako siłą produkcyjną, którą kapitał nieustannie stara się ujarzmić poprzez zarządzanie codzienną biopolityką” (Stark, Crawford, 2015, s. 2). Jak wiadomo z prac Hochschild (1983), praca emocjonalna, a współcześnie także praca afektywna, odnosi się do sposobów, w jakie aktywność afektywna w miejscu pracy i szerzej w sferze rynkowej stała się centralna dla funkcjonowania systemów ekonomicznych. Podobnie jak dane osobowe zostały skomodyfikowane, tak twarz emoji została nazwana „symbolem kapitalizmu” (BBC Radio and Wise Buddha Creative, 2013, cyt. za: Stark, Crawford, 2015). Autorzy ci argumentują, że platformy takie jak Facebook wykorzystują emoji i emocjonalne sygnalizatory w projektowaniu doświadczenia użytkownika w celu dalszej komodyfikacji i monetyzacji społecznego afektu (Stark, Crawford, 2015, s. 8). Opisują przy tym dość ponurą historię wzrostu znaczenia emoji, która została naznaczona pozwami sądowymi i przetworzona przez kapitalizm aż do wyczerpania zasobów znaczenia.

Wzorce użycia emoji między przyjaciółmi i partnerami mogą z czasem stać się abstrakcyjne i trudne do rozszyfrowania lub zdegenerować się do formy pro forma – zupełnie podstawowej. W najlepszym wypadku w wymianie takich symboli jak tęcza czy serce pojawia się osobisty podtekst, zawierający wiele warstw niewypowiedzianych znaczeń, których analitycy wywiadu czy algorytmy uczenia maszynowego nie są w stanie uchwycić. Niemniej ta złożoność nie powstrzymuje instytucji przed próbami jej komercjalizacji i wykorzystywania społecznej afektywności na różne sposoby (Stark, Crawford, 2015, s. 6). Testy A/B to zasadniczo eksperymenty, w których użytkownikom losowo prezentuje się różne warianty strony, aby za pomocą analizy statystycznej określić, która wersja skuteczniej realizuje zamierzony cel konwersji, taki jak polubienia, komentarze czy udostępnienia. Marketery szeroko twierdzą, że emoji zwiększają zaangażowanie użytkowników (Van Grove, 2013). Platformy takie jak Facebook, Twitter czy inne media społecznościowe integrują emoji z narzędziami analityki sentymentu. Serwis Emojitracker.com ukazuje zaskakującą ilość i różnorodność użycia emoji w czasie rzeczywistym na Twitterze, co czyni te dane cennymi dla cyfrowego kapitalizmu. Przykładowo, firma Hubspot wskazuje, że ich badania dowo-

dzą, iż niektóre emoji w szczególny sposób zwiększają liczbę kliknięć, komentarzy i udostępnień.

Co ciekawe, Hubspot zauważa, że najpopularniejsze emoji w badanym okresie (w okolicach Walentynek) nie należą do żółtych twarzy, co sugeruje, że mniej popularne emoji mogą być korzystniejsze w kampaniach marketingowych. Ich analiza jest jednak powierzchowna – nie są w stanie odpowiedzieć np. na pytanie, dlaczego emoji wiśni cieszy się tak dużą popularnością. Na blogu Hubspota można pobrać zestaw testów A/B, które mają na celu udowodnienie, że emoji zwiększają ruch oraz które z nich są najskuteczniejsze. Dla przykładu, emoji w poście na Facebooku rzekomo zwiększają liczbę polubień o 57%. Najczęściej używanym emoji na Twitterze przez trzy kolejne lata był „face with tears of joy”, jednak według Google Trends, w 2016 roku został on wyprzedzony przez czerwone serce, które obecnie dominuje jako najczęściej używany znak. 75% używanych emoji ma pozytywny wydźwięk, ale negatywne emoji pojawiają się częściej wieczorem, osiągając szczyt około godziny 21:00 (brandwatch.com). Niektóre z popularnych emoji są zaskakujące – na przykład ośmiornica zajęła pierwsze miejsce w rankingu klikalności, co autorzy uznają za przejaw „losowości przyciągającej uwagę odbiorcy”. Taka nieprzewidywalność bywa frustrująca dla korporacji, które dążą do pewności. Przewidywanie na podstawie analityki nadwyżki behawioralnej w przypadku symboli o płynnym znaczeniu nie jest proste.

Symbole nie są dane raz na zawsze – ich znaczenia się zmieniają. Narzędzia takie jak Google Trends pokazują, że różne emoji mają różną wartość afektywną w zależności od kraju, co świadczy o ich kulturowej specyfice. Przykładowo, z okazji Światowego Dnia Emoji 17 lipca 2018 roku Facebook opublikował analizę użycia emoji na swojej platformie. Najrzadziej używane emoji to m.in. „Man in Suit Levitating: Dark Skin Tone”, który według Google Trends nie posiada wystarczającej ilości danych do analizy. Z kolei inne wersje kolorystyczne tej samej postaci były częściej używane w USA i Indiach. Dane te wskazują na potencjalne uprzedzenia afektywne powiązane z rasą i reprezentacją. Joe Veix w *Newsweeku* (2016) opisał historię emoji Levitating Businessman, wywodzącego się z fontu Webdings Microsoftu z lat 90-tych – był to obrazek „rude boya” z okładki płyty zespołu *The Specials*, który w Unicode został przekształcony w biznesmena unoszącego się w powietrzu. Vincent Connare, twórca czcionki, uznał, że dwutonowy człowiek pasuje do czarno-białego stylu fontu i zainspirował się piosenką „Jump” ze wspomnianego albumu. Ironią jest to, że oryginalna postać stworzona przez Jerry’ego Dammersa była wzorowana na czarnoskórym muzyku reggae Peterze Toshu, członku *The Wailers*. Emojis są więc często kulturowo adoptowane i reprodukują względne znaczenia. Jak każda forma języka i konstrukcji dyskursywnych, mogą one sprzyjać tworzeniu baniek społecznych, w ramach których ich znaczenia ulegają polityzacji i podziałom. Dlatego ich użycie i wartość mogą ujawniać istotne informacje socjologiczne oraz psycho-

społeczne afektywności, które powinny być przedmiotem analizy ze strony nauk społecznych. W ramach analizy sentymentu i emocji w tweetach, wykorzystywano zestaw ikon emotek, emoji i ideogramów, rozwijany m.in. w Polsce (Wolney), które mają znaczenia kulturowo relatywne. Co uderzające, tak jak większość analiz sentymentu, są one redukowane do trzech głównych kategorii: pozytywny, negatywny i neutralny. Tego rodzaju klasyfikacja radykalnie upraszcza potencjalne znaczenia, ale w celach marketingowych – np. przy analizie hashtagów związanych z partią polityczną lub produktem – może być użyteczna. Jednocześnie brakuje w tych analizach tzw. afektywnego niuansu, który odzwierciedlałby złożoność i zmienność emocji w codziennej komunikacji cyfrowej.

PROCESUALNOŚĆ EMOCJI A ALGORYTMICZNA STANDARYZACJA AFEKTU: EMOJI, SENTYMENT I CYFROWA INFRASTRUKTURA EMOCJONALNA

Emotikony, emoji i ideogramy są wykorzystywane nie tylko jako narzędzia ekspresji emocjonalnej, ale również jako komponenty infrastruktury analitycznej, która przekształca afektywność w dane. Przykładowo, system Wolney – opracowany w Polsce – analizuje emocje w tweetach za pomocą zestawu symboli, przypisując im jedną z trzech wartości: pozytywną, negatywną lub neutralną. Choć narzędzie to może być użyteczne z punktu widzenia marketingu lub analizy dyskursu politycznego, redukuje ono wieloznaczność i złożoność afektywną do schematu binarnego, co skutkuje potencjalnym zafałszowaniem rzeczywistych doświadczeń emocjonalnych (Ellis, Tucker, 2020, s. 38). Dodatkowo, jak podkreślają Ellis i Tucker, brak „pozytywnych” symboli w niektórych zestawach sugeruje uprzedzenie systemowe, które może faworyzować negatywne lub neutralne odczytania wypowiedzi użytkowników. Jest to szczególnie niepokojące w kontekście algorytmicznej automatyzacji emocji, w której subtelność zostaje zredukowana do prostych wskaźników – kliknięć, reakcji, uśmiechów lub złości. Jak zauważają Novak i współpracownicy (2015), choć większość emoji jest interpretowana jako pozytywna, ich znaczenie jest głęboko kontekstualne. Na przykład emoji 🍆, które w niektórych kręgach jest kodem seksualnym, nie może być jednoznacznie sklasyfikowane bez znajomości intencji i kontekstu kulturowego. Sentiment analysis – w obecnym kształcie – nie jest w stanie wychwycić takich niuansów semantycznych, co podważa jej skuteczność i epistemologiczną wiarygodność. Podobnie, cytowana przez autorów fraza z Run D.M.C. – „Not bad meaning bad but bad meaning good!” – doskonale ilustruje niestabilność i wieloznaczność języka afektów, której nie sposób w pełni uchwycić za pomocą obecnych technologii analizy emocji. W tym sensie, znaczenia

emoji nie są statyczne, lecz procesualne i negocjowalne – są wynikiem relacji, intymności, humoru i lokalnych kontekstów.

Badacze Wiseman i Gould (2018) podają przykład pary, która wysyła sobie emoji pizzy jako wyraz miłości. To oznacza, że znaczenie emoji nie jest tylko efektem ogólnych kodów kulturowych, lecz także produktem subiektywnych mikro kultur emocjonalnych. Emoji stają się wówczas prywatnym kodem – rodzajem szyfru, który wymyka się analizie algorytmicznej. To właśnie ta niemożliwość pełnego uchwycenia sensu przez technologie analityczne staje się punktem napięcia między afektem jako potencjałem, a afektem jako produktem (Stark, Crawford, 2015). W cyfrowym kapitalizmie, jak zauważa Shoshana Zuboff (2019), każdy mikrogest użytkownika może stać się behawioralnym nadmiarem (behavioural surplus), który zasila modele predykcyjne korporacji. Podsumowując te rozważanie, należy stwierdzić, że emoji – choć z pozoru trywialne – odgrywają kluczową rolę w biopolityce emocji. Utrwalają, upraszczają i kategoryzują afekt, jednocześnie umożliwiając i ograniczając jego wyrażenie. W ich użyciu i analizie ujawniają się napięcia między:

- 1) emocją a jej reprezentacją;
- 2) prywatnością a ekstrakcją danych;
- 3) afektem jako ekspresją a afektem jako narzędziem władzy.

Badania Fan i in. (2018) analizujące ponad 70 tys. użytkowników Twittera wykazały, że akt etykietowania emocji („czuję złość”, „czuję smutek”) prowadził do regulacyjnego efektu: szybki spadek intensywności afektu w przypadku emocji negatywnych, wolniejszy w przypadku pozytywnych. Autorzy uznają to za dowód na efekt terapeutyczny etykietowania emocji. Jednak Ellis i Tucker krytycznie odnoszą się do tej interpretacji, podkreślając, że utożsamianie etykietowania emocji z ich ekspresją zakłada istnienie „wewnętrznej treści”, która może zostać „uwolniona” – co wpisuje się w hydrauliczny model emocji znany z psychologii poznawczej. Jak wskazują Ellis i Cromby (2009, 2011), taki model nie uwzględnia społecznego, performatywnego i relacyjnego charakteru afektu.

Zgodnie z Liu (2015), analiza sentymentu, choć rozwija się w kierunku większej złożoności (np. system EARL z 48 typami emocji), nadal sprowadza emocjonalność do wartościowania pozytywnego, negatywnego lub neutralnego. Jest to uderzająco uproszczony model, nieuwzględniający afektywnej płynności, kontekstu kulturowego czy subtelnej ekspresji emocjonalnej. Przykładowo, w systemach takich jak LIWC, emocja traktowana jest jako obiektywny wskaźnik możliwy do wyodrębnienia z tekstu – co ignoruje rolę ironii, subwersji i emocjonalnego kodowania znanego z mediów społecznościowych (Fan et al., 2018). Sentiment analysis, jak zauważają autorzy, nie radzi sobie z emocjonalnym tonem, dwuznacznością i lokalną semantyką afektywną.

Autorzy podkreślają, że platformy społecznościowe nie są neutralne – ich architektura została zaprojektowana nie tylko po to, by zaspokajać potrzeby, ale by je wytwarzać. Algorytmy emocjonalne, A/B testy, analiza sentymentu i ograniczone zestawy reakcji (jak pięć emoji Facebooka) są narzędziami kapitalizacji afektu (Zuboff, 2019; Bucher, 2018). Podążając za teorią programmed sociality (Bucher, 2018), Ellis i Tucker pokazują, że afekty nie są jedynie wyrażane w mediach społecznościowych – są generowane i modelowane. To tworzy nowe formy emocjonalnej pracy (Hochschild, 1983; Ahmed, 2004), w których użytkownicy stają się afektywnymi pracownikami platformy, produkującymi emocjonalne dane jako wartość ekonomiczną. Podsumowując, autorzy wskazują na fundamentalne napięcie: afekt jako towar kontra afekt jako relacyjna, kulturowa i polityczna siła. Redukcja emocjonalności do binarnego systemu (pozytywne/negatywne) służy logice ekstrakcji danych, a nie zrozumieniu człowieka. Ellis i Tucker apelują o badania emocji dla celów ludzkich, nie kapitalistycznych – dążące do transindywidualności (Stiegler, 2010) i komunitaryzmu emocjonalnego, a nie tylko do precyzyjniejszego targetowania reklam.

Praca emocjonalna (emotional labour), pojęcie wprowadzone przez Hochschild (1983), opisuje proces zarządzania własnymi emocjami w celu sprostania wymogom zewnętrznym, najczęściej zawodowym. W kontekście mediów społecznościowych ta praca przybiera nową formę: użytkownicy stają się nieświadomymi pracownikami platform, których emocje i reakcje są wykorzystywane jako surowiec do generowania danych o wysokiej wartości ekonomicznej (Ahmed, 2004; Bucher, 2018). Ellis i Tucker (2020) wskazują, że w erze cyfrowej praca emocjonalna ulega programowalnej automatyzacji. Algorytmy analizujące sentyment i afekt kształtują zachowania użytkowników poprzez określone bodźce, takie jak zestawy reakcji emoji (Stark i Crawford, 2015) lub zmienne wyświetlanie treści w oparciu o analizę afektywną. W efekcie afekt nie jest już tylko subiektywnym przeżyciem, ale integralnym elementem mechanizmu ekonomicznego. Systemy analizy sentymentu, takie jak LIWC czy EARL, upraszczają bogactwo ludzkich emocji do kilku kategorii (pozytywne, negatywne, neutralne), co powoduje istotne zniekształcenia interpretacyjne (Liu, 2015). Takie podejście ignoruje społeczno-kulturowy kontekst emocji, ich performatywną naturę oraz afektywną ambiwalencję. Przykładowo, ironia i subwersja w komunikacji internetowej mogą całkowicie zmieniać znaczenie pozornie negatywnej wypowiedzi. Systemy sentymentu nie radzą sobie również z lokalnymi wariantami językowymi i złożonymi formami wyrazu emocji, co prowadzi do błędnych klasyfikacji i utraty niuansów (Fan et al., 2018; Nash, 2016).

Bucher (2018) opisuje zjawisko programmed sociality, w którym algorytmy nie tylko odzwierciedlają, ale i aktywnie kształtują zachowania użytkowników. Ellis i Tucker (2020) rozwijają tę ideę, podkreślając, że media społecznościowe generują i modulują emocje, integrując je z mechanizmami

ekonomicznymi kapitalizmu nadzoru (Zuboff, 2019). Użytkownicy, często nieświadomi tego procesu, stają się „afektywnymi robotnikami” (emotional workers), produkującymi dane, które platformy wykorzystują do personalizacji reklam, zwiększania zaangażowania i generowania zysków. Ta ekonomia afektu opiera się na ciągłej produkcji, monitoringu i optymalizacji emocjonalnej aktywności w sieci. Zakończenie refleksji Ellis i Tuckera (2020) zwraca uwagę na konieczność wyjścia poza kapitalistyczną logikę eksploatacji afektu. Autorzy odwołują się do koncepcji transindywidualności Stieglera (2010), czyli kolektywnej podmiotowości powstającej w relacjach społecznych i kulturowych, w której emocje mają znaczenie nie tylko indywidualne, ale także polityczne i etyczne. Apelują o rozwój badań i praktyk społecznych, które uwolnią emocje z pułapki uproszczeń analizy sentymentu i komercyjnego wykorzystywania. Wskazują na potrzebę komunitaryzmu emocjonalnego, gdzie afekt służy budowaniu więzi, oporu i solidarności, a nie jedynie targetowaniu reklam. Analiza sentymentu i praktyka etykietowania emocji w mediach społecznościowych stanowią obecnie kluczowe narzędzia technologicznej ingerencji w afektywność użytkowników. Pomimo pewnych terapeutycznych interpretacji tych zjawisk, krytyczna refleksja Ellis i Tuckera (2020) ujawnia ich głębszy wymiar ekonomiczny i społeczny, gdzie emocje są redukowane do danych, zarządzane i eksploatowane. Konieczne jest poszerzenie perspektywy badawczej, uwzględniającej społeczne, kulturowe i performatywne aspekty afektu, aby móc przeciwstawić się upraszczającej i instrumentalnej logice cyfrowego kapitalizmu emocjonalnego.

W dobie mediów społecznościowych emocje stają się integralnym składnikiem afektywnej infrastruktury cyfrowej. Tak jak Nash (2016) podkreśla, codzienne interakcje online nie tylko wyrażają emocje, ale także wchodzą w skład złożonych systemów danych, które podlegają algorytmicznej obróbce i rekontekstualizacji. Platformy takie jak Facebook, Twitter czy Instagram projektują swoje interfejsy tak, by emocje użytkowników były łatwo mierzalne i kwantyfikowalne. Przykładem są reakcje emoji czy automatyczne tagowanie afektu, które umożliwiają gromadzenie ogromnych ilości danych emocjonalnych, poddawanych dalszej analizie w celu optymalizacji doświadczenia użytkownika i zwiększenia zysków korporacyjnych (Bucher, 2018; Zuboff, 2019). Ta infrastruktura jednak, jak ostrzegają Ellis i Tucker (2020), niesie ze sobą ryzyko uproszczenia i instrumentalizacji afektu, redukując bogate, społeczne i kulturowe znaczenia emocji do jednostek analitycznych pozbawionych kontekstu.

Ellis i Cromby (2009, 2011) zwracają uwagę, że emocje nie są statycznymi treściami wewnętrznymi, ale dynamicznymi procesami społecznymi i performatywnymi. Oznacza to, że emocje kształtują się i manifestują w relacjach międzyludzkich, w określonych kontekstach kulturowych i historycznych. Tymczasem technologie cyfrowe, poprzez swoją architekturę i logikę działania, często ignorują ten wymiar, traktując emocje jako obiektywne

dane do zbierania i analizy. To powoduje, że kontekst, intencje czy historyczne uwarunkowania emocji zostają pominięte, co może prowadzić do błędnych interpretacji i utraty złożoności afektywnej komunikacji (Liu, 2015; Nash, 2016). Algorytmy wykorzystywane w mediach społecznościowych nie tylko rejestrują emocje, ale również kształtują normy afektywne, czyli społeczne oczekiwania dotyczące tego, jakie emocje są odpowiednie, a jakie nie (Bucher, 2018). Przykładowo, ograniczony repertuar reakcji emoji promuje uproszczone modele emocjonalne, które mogą wpływać na sposób, w jaki użytkownicy wyrażają swoje uczucia. Stark i Crawford (2015) wskazują, że emoji często reprodukuje konserwatywne i stereotypowe wzorce ekspresji emocji, co ogranicza afektywną różnorodność i może prowadzić do marginalizacji mniej typowych form ekspresji.

Mimo tych ograniczeń, media społecznościowe stanowią także przestrzeń dla oporu afektywnego – praktyk, które podważają zredukowane modele emocji i wyrażają alternatywne formy afektywności. Ellis i Tucker (2020) podkreślają, że cyfrowy kapitalizm emocjonalny może być jednocześnie przestrzenią solidarności i współdziałania, czyli tzw. komunitaryzmu emocjonalnego. Transindywidualność (Stiegler, 2010) to koncepcja, która wskazuje na możliwość budowania zbiorowych podmiotowości, w których emocje stają się narzędziem wspólnego działania i politycznej zmiany, a nie jedynie surowcem do komercyjnej eksploatacji. Badania nad emocjami w mediach społecznościowych wymagają refleksji etycznej, zwłaszcza w kontekście wykorzystywania danych emocjonalnych bez świadomej zgody użytkowników (Fuchs, 2017; Zuboff, 2019). Problem ten dotyczy zarówno ochrony prywatności, jak i zagrożeń związanych z manipulacją afektywną, która może wpływać na wybory polityczne, konsumenckie czy społeczne. Ellis i Tucker (2020) postulują, by badania nad emocjami koncentrowały się na ochronie dobrostanu ludzi i wzmacnianiu ich autonomii, a nie na maksymalizacji zysków korporacji czy zwiększaniu efektywności marketingu. W obliczu wyzwań współczesnej ekonomii afektu, konieczne jest poszukiwanie alternatywnych modeli afektywności, które uwzględniają złożoność emocji, ich społeczne znaczenie i potencjał wyzwalający. Prace filozofów takich jak Stiegler (2010) oraz badaczy afektywnych praktyk społecznych (Ahmed, 2004; Nash, 2016) wskazują na możliwość tworzenia cyfrowych przestrzeni opartych na solidarności i współpracy, które mogłyby stanowić przeciwwagę dla dominującego kapitalizmu nadzoru.

BIBLIOGRAFIA

- A S. Ahmed, *The Cultural Politics of Emotion*, Edinburgh University Press, Edinburgh 2004.
- C. Beasley, D. Mason, *Emotion, Subjectivity, and the Mediation of Sociality: Affect and Emotional Labour Reconsidered*, *Emotion, Space and Society*, 14, 2015, s. 1–8.
- T. Bucher, *If... Then: Algorithmic Power and Politics*, Oxford University Press, Oxford 2018.
- P. Chen, M. Potenza, *Psychographic profiling and voter targeting: Efficacy and limitations*, *Journal of Political Marketing*, 17(2), 2018, s. 145–160.
- D. Ellis, *Data and identity in social media*, *Digital Culture & Society*, 4(1), 2018, s. 14–30.
- D. Ellis, J. Cromby, *Emotional Expression and Disclosure: A Critical Review*, *Psychology and Psychotherapy: Theory, Research and Practice*, 82(3), 2009, s. 337–353.
- D. Ellis, J. Cromby, *Social Psychology of Emotion*, Sage, London 2011.
- D. Ellis, I. Tucker, *Emotion in the Digital Age: Technologies, Data and Psychosocial Life*, Routledge, London 2020.
- D. Ellis, I. Tucker, *Emotion in the Digital Age*, Palgrave Macmillan, London 2021.
- R. Fan, J. Zhao, Y. Chen, J. Sun, K. Xu, *Anger Is More Influential Than Joy: Sentiment Correlation in Weibo*, *PLOS One*, 13(3), 2018, e0193503.
- C. Fuchs, *Social Media: A Critical Introduction*, Sage, London 2017.
- E. Graham-Harrison, C. Cadwalladr, *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*, *The Guardian*, 17 March 2018.
- R. Hochschild, *The Managed Heart: Commercialization of Human Feeling*, University of California Press, Berkeley 1983.
- D. I. Kramer, J. E. Guillory, J. T. Hancock, *Experimental evidence of massive-scale emotional contagion through social networks*, *Proceedings of the National Academy of Sciences*, 111(24), 2014, s. 8788–8790.
- Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, Cambridge University Press, Cambridge 2015.
- W. Mischel, *Personality and Assessment*, Wiley, New York 1968.
- E. Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism*, PublicAffairs, New York 2013.
- K. Nash, *Affective Encounters: Everyday Life and the Digital*, Palgrave Macmillan, London 2016.
- M. Nash, *Digital Capitalism and Affective Labour*, Routledge, London 2016.
- J. W. Pennebaker, *Opening Up: The Healing Power of Expressing Emotions*, Guilford Press, New York 1997.
- J. Pridmore, *Consumption and Identity in the Age of Social Media*, *Information, Communication & Society*, 15(4), 2012, s. 574–596.
- H. Rheingold, *The Virtual Community: Homesteading on the Electronic Frontier*, Addison-Wesley, Reading, MA 1994.
- L. Stark, K. Crawford, *The Conservatism of Emoji*, *Social Media + Society*, 1(2), 2015.
- Stiegler, *Taking Care of Youth and the Generations*, Stanford University Press, Stanford 2010.
- K. Taggart, *Did Cambridge Analytica really influence the 2016 US election?*, *The Intercept*, 10 November 2017.
- Tucker, D. Ellis, *Emotion in the Digital Age*, Palgrave Macmillan, London 2021.
- N. Whitehead, *Process and Reality*, D. R. Griffin, D. W. Sherburne (red.), Corrected edition, Free Press, New York 2011.
- R. Williams, *The Long Revolution*, Chatto & Windus, London 1961.
- W. Youyou, M. Kosinski, D. Stillwell, *Computer-based personality judgments are more accurate than those made by humans*, *Proceedings of the National Academy of Sciences*, 112(4), 2015, s. 1036–1040.
- S. Zuboff, *The Age of Surveillance Capitalism*, PublicAffairs, New York 2019.

**EMOTIONAL MANIPULATION AND AFFECTIVE DATA IN THE
DIGITAL AGE**

ABSTRACT

This article offers a critical analysis of emotional manipulation in social media, framing it as a constitutive element of the broader infrastructure of digital affective capitalism. It opens with empirical cases of interference in users' emotional states—the Facebook emotional contagion experiment of 2014 and the activities of Cambridge Analytica—which reveal the capacity of digital platforms to modulate moods, decisions, and political behaviors through algorithms, personal data, and psychographic microtargeting. At the same time, the article argues that narratives of the “omnipotence” of these technologies are often exaggerated, as models of personality and emotional prediction remain epistemologically fragile, reductive, and highly context-dependent.

The analysis then shifts from spectacular scandals to everyday affective practices, such as the use of emoji, emoticons, and platform-specific reaction systems. These are interpreted as tools of emotional standardization that simultaneously enable affective expression and reduce its complexity, rendering it susceptible to aggregation, sentiment analysis, and commodification. Drawing on Whitehead's process philosophy, Butler's theory of performativity, Hochschild's concept of emotional labor, and Zuboff's critique of surveillance capitalism, the article argues that emotions in social media are not merely expressed but are co-produced by platform architectures.

Emoji and sentiment analysis systems are thus presented as components of a biopolitics of emotion, in which affect becomes an economic resource and users function as affective workers for platforms. At the same time, the article highlights the limits of algorithmic approaches to emotion: its contextuality, ambiguity, and relational character resist full automation. In conclusion, an alternative research perspective is proposed, inspired by concepts of transindividuality and emotional communitarianism, which conceptualizes affect not as a commodity but as a relational, political, and ethical resource capable of fostering solidarity and resistance to the instrumental logic of digital capitalism.

Keywords: emotional manipulation; social media; emoji; sentiment analysis; affect; emotional labor; surveillance capitalism; personal data; algorithms; psychographic microtargeting; transindividuality; biopolitics of emotion.

O AUTORZE – adiunkt, dr hab., pracuje w Instytucie Filozofii UMCS Lublin, dziedziny filozoficzne: filozofia fikcji, filozofia emocji, ontologia.

E-mail: jacek.gurczynski@mail.umcs.pl

Sławomir Czetwertyński,
Jakub Marcinkowski

O HEURYSTYKACH BADAWCZYCH W NAUCE INSTYTUCJONALNEJ W PERSPEKTYWIE NAUK EKONOMICZNYCH

<https://doi.org/10.37240/FiN.2025.13.06>

STRESZCZENIE

W artykule zajęto się problematyką dominacji metody izolacji w naukach ekonomicznych, w kontekście uwarunkowań instytucjonalnych i behawioralnych. Autorzy stawiają hipotezę, że naukowcy posługują się heurystykami badawczymi w ramach nauki instytucjonalnej, rozumianej za Kuhnem jako okres konsensusu paradygmatycznego. Metoda izolacji, jako forma redukcjonizmu wywodząca się z Weberowskiej koncepcji typu idealnego, stanowi nie tylko narzędzie poznawcze upraszczające złożone zjawiska społeczno-gospodarcze, ale również bezpieczną strategię kształtowania własnej kariery akademickiej. W ujęciu behawioralnym badacz, jako podmiot poznawczo ograniczony, skłonny jest do stosowania sprawdzonych metod, które można porównać do heurystyk decyzyjnych znanych z ekonomii behawioralnej. Presja instytucjonalna wyrażająca się maksymą „publikuj albo gin” oraz biologiczny imperatyw gospodarowania budżetem energetycznym sprzyjają wykorzystywaniu utartych schematów badawczych. Artykuł wykazuje, że wybory metodologiczne ekonomistów są efektem interakcji między tradycją badawczą i ograniczeniami poznawczymi a wymogami kariery akademickiej.

Słowa kluczowe: nauka instytucjonalna, metoda izolacji, heurystyki badawcze, metodologia nauk ekonomicznych.

1. WPROWADZENIE

Maks Weber (1984, s. 70) jeszcze na przełomie XIX i XX wieku zwracał uwagę na fakt, że to co zwiemy „nauką” jest *de facto* obowiązującym standardem przyjętym w kulturze Zachodu. Jego uwaga na temat konstruktywizmu w nauce, który za jego czasów nie był jeszcze tak silnie akcentowany jak obecnie, ukazuje prymat konkretnej kultury w światowym życiu academic-

kim. Uwagę tę Weber popęlnił co prawda w kontekście rozważań nad historycznymi uwarunkowaniami rozwoju cywilizacji europejskiej, jednak istotny aspekt związany jest z samą naturą relacji w poszczególnych grupach społecznych. Społeczność akademicka, dla której nauka jako wytwór kultury jest centralnym elementem struktury wartości, podlegać będzie tym samym co do istoty zjawiskom, co każda inna grupa społeczna skupiona wokół własnych wytworów kultury. I tak samo jak w przypadku innych grup podlegać będzie kontroli społecznej (Berger, 1963) z ostracyzmem włącznie.

Historia myśli naukowej dostarcza szeregu przypadków, gdy badacz musiał zmagać się z ogólnym problemem wykluczenia lub marginalizacji z życia akademickiego ze względu na poglądy, które nie były akceptowane w jego środowisku. Przykładem takim jest chociażby Ludwigo Pasteur, którego tezy dotyczące mikroorganizmów były w sprzeczności z powszechną za jego czasów teorią o spontanicznym powstawaniu życia. Podobnie George A. Akerlof musiał zmagać się z trudnościami publikacyjnymi dotyczącymi jego najśłynniejszej pracy na temat asymetrii informacji (Akerlof, 2003). Z kolei biografowie Alberta Einsteina wskazują, że jego dysertacja doktorska na temat nowych metod wyznaczania rozmiarów molekuł była osadzona w istniejącej tradycji badawczej i nie była bezpośrednio związana z jego późniejszymi rewolucyjnymi odkryciami. Co więcej analiza tematyki pracy oraz etapy jej powstawania wskazuje się jako argumenty na rzecz ograniczeń instytucjonalnych występujących w niemieckojęzycznych kręgach akademickich (Caprice et al., 2008).

Trudności z wprowadzeniem do nauki poglądów niezgodnych z ogólnie przyjętym konsensusem są problemem dobrze znanym. W zasadzie od chwili upowszechnienia się teorii rewolucji naukowej Thomasa Kuhna (1962) jasnym jest, że wymiar socjologiczny paradygmatów naukowych tworzy ramy funkcjonowania w społeczności akademickiej (por. Motycka, 1980). Okres konsensusu polegający na utrzymywaniu się *status quo* wśród akademików, który Kuhn określa mianem „nauki normalnej”, obfituje w rozwiązywanie łamigłówek intelektualnych, które są zgodne z przyjętymi standardami. Łamigłówki te w żadnym razie nie mają zaburzyć owego *status quo*, co rzutuje zarówno na tematykę badań, jak i przyjęte metody.

W polskim przekładzie pracy Kuhna z 1968 roku Stefan Amsterdamski zaproponował tłumaczenie dynamiczne, w którym zamiast „nauki normalnej” zastosował określenie „nauki instytucjonalnej” (por. Kuhn, 1968; 2009). Jest ono w tym przypadku zdecydowanie trafniejsze niż *verbum pro verbo*, gdyż wskazuje silny nacisk ram normatywnych, które towarzyszą życiu akademickiemu. W tym przypadku przekład *sensum pro sensum* ukazuje fakt, że kariera akademicka nierozzerwalnie łączy się z formalnymi i nieformalnymi wymogami, które badacz musi spełnić na drodze do awansu w hierarchii zawodowej. I chociaż ideałem nauki pozostaje poszukiwanie

prawdy, to nie sposób ustrzec się przed pozornym celem nauki jakim jest zdobywanie stopni i tytułów naukowych.

Skłania to do postawienia ogólnej tezy głoszącej, że badacze stosują uznane w środowisku akademickim metody rozwiązywania problemów badawczych w celu rozwoju własnej kariery akademickiej. Tak postawiona teza jest jednak ułomna, gdyż nie sposób wykazać, że zjawisko to ma charakter permanentny. Jest to raczej zjawisko powszechne o pewnym natężeniu, które z pewnością będzie ulegało zmianom. Z kolei stawianie tezy dotyczącej pewnej tendencji w środowisku akademickim nosiłoby znamiona truizmu. Wiadomo, że akademicy wykorzystują ogólnie przyjęte problemy badawcze i metody w swojej pracy zawodowej, ale nie ulega również wątpliwości, że dokonują odkryć naukowych, zarówno w obrębie przyjętego konsensusu, jak i wychodząc mniej lub bardziej poza ramy funkcjonującego w środowisku akademickim *status quo*. Z tego względu w niniejszym artykule zdecydowano się zastosować ujęcie oparte na czynnikach psychologicznych osadzonych w ramach instytucjonalnych. Oznacza to wykorzystanie analogii do dorobku ekonomii behawioralnej, w której powszechnie wykorzystuje się pojęcie *heurystyk*, rozumianych jako ograniczoną liczbę uproszczonych zasad wnioskowania, które redukują zadania złożone, takie jak ocena prawdopodobieństwa czy prognozowanie, do prostych operacji oceniających zjawisko (Tversky, Kahneman, 1974). Uproszczone operacje oceny danych zjawisk charakteryzują się wysokim stopniem efektywności (w kontekście relacji kosztów ich wykorzystania i korzyści jakie niosą ze sobą) w procesie decyzyjnym, aczkolwiek w niektórych przypadkach narażone są one na błędy.

W badaniach z zakresu socjologii nauki i metodologii nauki zastosowanie ujęcia opartego na heurystykach oznacza przyjęcie paraleli między ogólnymi wyborami ludzkimi a wyborami metodologicznymi. W takim kontekście zaproponować można hipotezę głoszącą, że *naukowcy posługują się w badaniach naukowych heurystykami w ramach nauki instytucjonalnej*. Wykazanie poprawności tak postawionej hipotezy implikuje pytanie o samą heurystykę, a zatem metodę doboru problemu badawczego oraz metodę jego rozwiązania. W niniejszym artykule rozważania zawężone będą do obszaru nauk ekonomicznych i powszechności głównego podejścia badawczego występującego w ortodoksji ekonomicznej, czyli do izolacji. Tym samym pytanie badawcze niniejszego artykułu przyjmuje postać: *dlaczego ekonomiści stosują metodę izolacji do rozwiązywania problemów ekonomicznych?*

Tak postawiony cel oraz pytanie badawcze determinują strukturę artykułu. W pierwszej części artykułu zaprezentowane będzie ujęcie instytucjonalne uprawiania nauki, które jest dobrze rozpoznane przez filozofów nauki i historyków nauki. W kolejnej części artykułu rozważania poszerzone zostaną o perspektywę behawioralną, a zatem czynniki psychologiczne właściwe badaczom. Ostatnia część artykułu będzie formą syntezy obu podejść, a za-

tem wykazania jak naukowiec jako podmiot psychospołeczny, działa w warunkach środowiska akademickiego.

2. UJĘCIE INSTYTUCJONALNE UPRAWIANIA NAUKI

W niniejszym artykule przyjęto za Kuhnem (1962) podział etapów rozwoju nauki na okres *nauki normalnej* i moment *rewolucji naukowej*. Przy czym nauka normalna określana będzie za tłumaczeniem Stefana Amsterdamskiego jako *nauka instytucjonalna* jak to zostało nakreślone we wprowadzeniu. Wynika to z tego, że w świetle prowadzonych tu rozważań istotny jest nie tyle ideał nauki poszukującej prawdy, co aspekt kariery akademickiej wiążący się nieuchronnie z koniecznością spełniania wymogów formalnych i publikacji prac rygorowych. W ramach uprawiania nauki instytucjonalnej *nie* należy zakładać, że praca naukowa będzie źródłem *tez* w rozumieniu Arystotelesa (1978), a zatem takich które stoją w sprzeczności do ogólnie przyjętych poglądów. Wynika to z samej definicji nauki instytucjonalnej, która nie przełamuje ogólnie panującego paradygmatu, więc z gruntu rzeczy tezy powstałe w jej ramach powinny być zgodne z panującym wśród akademików konsensusem – muszą być zgodne z pewną konwencją. Niemniej należy zaznaczyć, że jest to pewne rozmyślne uproszczenie, gdyż nawet w nauce instytucjonalnej tezy powinny być odkrywcz. I o ile jest to myślenie życzeniowe, to posiłkując się myślą Imre Lakatosa (1978), nawet zdegenerowane naukowe programy badawcze mogą stać się źródłem nowych zaskakujących tez. W artykule położono jednak akcent głównie na pracach przyczynkarskich, które nie zwiększają liczby anomalii mających być zalążkiem rewolucji naukowej. Przypadki takie potraktowano jako endemiczne „czarne łabędzie” w rozumieniu Nassima Taleba (2020), a zatem mało prawdopodobne i rzadkie zjawiska, o dużym wpływie. Z pewnością takie się zdarzają i prowadzą do poważnych wstrząsów w nauce, lecz nie należy traktować ich jako obligatoryjnego elementu prozaicznej kariery akademickiej, realizującej pozorny cel nauki w postaci zdobywania stopni i tytułów naukowych.

O ile filozofowie nauki i metodolodzy wskazują na negatywne aspekty trwania na etapie nauki instytucjonalnej, to jest to zjawisko nie do uniknięcia, a często jest jak najbardziej pożądane, gdyż chroni przed masowym powstawaniem teorii fałszywych. Wyjątkowo jaskrawym przykładem niebezpieczeństw związanych ze zbytnim entuzjazmem w wychodzeniu z utartych schematów jest pseudo naukowa kariera jednego z fizyków z Bell Laboratories, którego prace okazały się oszustwem i musiały być wycofane z poczytnych i uznanych periodyków o światowym zasięgu (por. Reich, 2009; Service, 2002). Inny przykład dotyczy swoistego imperializmu statystyki w ekonomii. W tym przypadku znane jest stanowisko Milтона Friedmana (1953), który wyraźnie miał obiekcje przed wykorzystywaniem regresji wielorakiej

jako odpowiedzi na wszystkie problemy analityczne ekonomii i tym samym trwania w jednej konwencji badawczej. Jego uwagi dotyczące ekonomii pozytywnej stały się zarzewiem dyskusji nad zbytnią ufnością w zdolności poznawcze metod opartych właśnie na regresji wielorakiej (Blaug, 1995), co przypomina w zasadzie powtarzalne ćwiczenia w ramach przyjętego przez badacza paradygmatu.

Próbując znaleźć równowagę pomiędzy racjonalnością w nauce (Sady, 2013) a odwagą przełamywania utartych schematów badawczych (Leibenstein, 1988), trudno nie narazić się na zarzut myślenia życzeniowego odnośnie prozaicznych trudności ludzi nauki. Zatem podejście do tematu z punktu widzenia ustalenia zasad realizacji pracy naukowej nieuchronnie obarczone będzie ujęciem normatywnym, w którym wartościowanie – czy nawet moralizatorstwo – jest kwestią fundamentalną. Jednakże można zaproponować podejście odmienne, które opiera się na ujęciu indywidualistycznym, w którym badacza traktuje się jako podmiot rozwiązujący problem w ramach zastanego środowiska (w tym przypadku zinstytucjonalizowanego).

Badacz, jak każdy człowiek, z natury posiada ograniczone zdolności poznawcze. Dowodzonego tego zarówno na płaszczyźnie filozofii krytycznej (Kant, 2012), błędów indukcji (Popper, 2010), czy zdolności kognitywnych człowieka (Pinker, 2021). Niezależnie od tego człowiek ma w swojej naturze również potrzebę *poznania* (Maslow, 1990), co w pewnym sensie skazuje go na ciągłe poszukiwanie prawdy, której odnaleźć nie może (Grobler, 2000). Badacz jako podmiot społeczności akademickiej, problem poznania przesuwają na płaszczyznę zawodową lub inaczej profesjonalną. Zatem jego codzienność akademicka powinna być związana z koniecznością poszukiwania i rozwiązywania problemów badawczych.

Zasadnicza będzie tu *motywacja* badacza w kontekście profesjonalizacji zawodu badacza. Może ona wynikać z (1) wyidealizowanych pobudek, w efekcie czego badacz dąży do zajmowania się wyłącznie „wielkimi” problemami, próbując przyjąć pozycję ponad środowiskiem instytucjonalnym. Może ona również wynikać z (2) pragmatycznych pobudek (by nie powiedzieć oportunistycznych), które skorelowane są z określoną pozycją społeczno-zawodową w środowisku naukowym.

Badacza można ujmować jako osobę rozdartą pomiędzy wielowiekową tradycją naukową a współczesnym mu środowiskiem instytucjonalnym. Badacz jest w pewnym sensie niewolnikiem własnej epoki, który w zależności od właściwych mu pobudek, albo pozostaje pod wpływem wyidealizowanego obrazu nauki, albo koncentruje się na formalnych wymogach kariery akademickiej. W praktyce należy sądzić, że zasygnalizowana tu dychotomia będzie zwykle przyjmowała stan pośredni. Te z kolei rzutuje na skalę i rodzaj podejmowanych przez badacza problemów, które uzależnia on od instytucjonalnego systemu nauki wymagającego od niego wypełnienia określonych zadań. Te z kolei są warunkiem otrzymywania pozytywnej oceny za jego pra-

cę akademicką, co pośrednio zmusza go do konieczności starania się o kolejne stopnie i tytuły naukowe.

Samą pracę badawczą można uprościć do dwóch elementów bazowych: (1) określenia problemu badawczego i (2) rozwiązania tego problemu. Oba te elementy można ujmować w kategoriach decyzyjnych, zarówno w zakresie wyboru tego problemu, jak i następnie doboru metody jego rozwiązania. W postaci wyidealizowanej stanowiącej *sine qua non* uprawiania nauki będzie to sposób postępowania metodyczny, neutralny ontologicznie. Przy czym, jako że badacze są tylko i aż ludźmi, okazuje się, że wskazane dwie kategorie decyzyjne bywają niezdeterminowane kierunkowo, czyli klasycznie najpierw rodzącego się problemu badawczego, a następnie dobierania do niego metody badawczej. Argumentem na rzecz zmiękczenia wspomnianego postępowania w uprawianiu nauki jest stanowisko analogicznie do nauk fizykalnych, iż pomiar może zmienić badane zjawisko. Ponadto w samych naukach społecznych zidentyfikowano ich własność dotyczącą stosowania mierników w stosunku do problemów społecznych. Donald Campbell (1979) sformułował zależność między skwantyfikowaniem wskaźników społecznych stosowanych w podejmowaniu wyborów a podatności na zniekształcanie tych wskaźników. Jego wnioski są dość oczywiste, by nie powiedzieć trywialne i sprawdzają się do tego, że jeżeli stosować jakąś parametryzację procesów społecznych, to następują działania jednostek, które wypaczają jej sens. I to w tym świetle w nauce instytucjonalnej, w kontekście procesu decyzyjnego dotyczącego wyboru problemu badawczego i metody badawczej, skutkować to będzie rekurencją między problemem a metodą. Oznacza to nic innego, jak to, że problem dobierany będzie w kontekście metody, a w skrajnym i można rzec absurdalnym przypadku to metoda zdeterminuje problem.

Trudno nawet mówić, że opisana powyżej zależność zmiękcząca ortodoksyjne postępowanie w uprawianiu nauki jest czymś nieprawidłowym, gdy naukę bada się w ujęciu procesu społecznego, a sami adepci nauki uwikłani są w pozorną konieczność zdobywania stopni i tytułów naukowych. Ludwig Fleck (1936) podkreśla, że ten podmiot epistemologiczny, który można sprowadzić do określenia „badacz”, jest wytworem symbolicznym, którego cechy są kontryfakcyjne. Dla takiego „badacza” źródłem poznania miałyby być najpierw obserwacja, a zatem doświadczenie empiryczne. Jednak jest to wyłącznie jeden z możliwych modeli badacza, który nie oddaje istoty funkcjonowania w warunkach nauki instytucjonalnej. Model, który byłby w niej zasadny, opierać się będzie na fakcie, że „badacz” jest zakorzeniony w wiedzy akademickiej, która jest mu właściwa, którą przyswoił w procesie kształcenia od poziomu szkół podstawowych po szkoły wyższe i w trakcie pracy badawczej. A zatem jego wybory nie są ahistoryczne i aspołeczne, lecz umiejscowione są w panującym konsensusie zarówno w czasie, jak i przestrzeni społecznej – czyli w pewnej tradycji badawczej. I w tym sensie znajomość metod badawczych akceptowanych przez reprezentantów danej szkoły naukowej

pośrednio determinuje dobór problemów badawczych, co jest właściwie emblematyczne dla nauki instytucjonalnej.

Nauki ekonomiczne, mimo silnego podkreślania swojej dominacji, czyli imperializmu ekonomicznego, wywodzą się z rozważań ogólniejszych. Stąd poszukuje się ich źródeł w dyskursie filozofów (Hardt, 2013) i myślicieli społecznych (Weber, 2002). Trudno szukać konkretnych źródeł myśli ekonomicznej, jako że trudno uznać nauki ekonomiczne jako twór jednolity – jej dyscypliny i subdyscypliny mogą się bowiem cechować interdyscyplinarnością. Nie można z resztą jednoznacznie wydzielić z ekonomii jednego ściśle przestrzeganego zbioru założeń. Sam Kuhn (1977, s. 222) pisał, że nauki społeczne nie wykształciły jednego paradygmatu (co dyskryminowało je w stosunku do nauk przyrodniczych i ścisłych), a George Ritzer (1975) pisał, że raczej należy traktować nauki społeczne jako multiparadygmatyczne. Niemniej, dość eklektycznej identyfikacji paradygmatów w naukach społecznych podjęli się Gibson Burrell i Gareth Morgan (1979), którzy w ramach kryteriów wyróżnili preferowaną orientację społeczną oraz założenia epistemologiczne dotyczące ideału nauki. I tak zidentyfikowali funkcjonalizm, interpretatywizm, radykalny strukturalizm i postmodernizm, które osadzili w powstałej macierzy z dychotomicznymi relacjami: regulacja *us* radykalna zmiana oraz obiektywizm *us* subiektywizm. W każdym razie, ten paradygmatyczny dyskurs sprowadza się to tego, że różne tradycje badawcze mogą rozwijać się symultanicznie w różnych przestrzeniach społecznych. Florian Znaniecki (1984) wyraźnie ukazał taki stan rzeczy opisując podział nauk na szkoły, co zdaje się trafnie obrazować tradycje badawcze danych grup badaczy.

Mimo, że nauki ekonomiczne wydają się dość spójne, to bliższe ich poznanie ukazuje różne podejścia do kwestii formułowania problemów i stosowania metod badawczych. Chociaż trudno określić gnoseologię (patrz. Motycka, 1980) poszczególnych szkół, czy też tradycji badawczych, to w ekonomii ortodoksyjnej (wystarczająco sformalizowanej by móc odtworzyć fundamentalne założenia) istotny wydaje się wpływ myśli metodologicznej Webera. Przejawia się ona w szeroko stosowanym w naukach ekonomicznych redukcjonizmie. Redukcjonizm, szczególnie przydatny w warunkach ograniczonych możliwości poznawczych, daje narzędzia do deformacji rzeczywistości do stanu, który pozwala mu na wyciągnięcie wniosków i formułowanie praw naukowych. Izolacja (Mäki, 1992), idealizacja (Nowak, 1980), abstrakcja (Lange, 1986), modelowanie (Arrow, Hahn, 1991) to jedynie niektóre przykłady opisujące co do idei ten sam zbiór metod poznawczych. Cytowani powyżej badacze podkreślają atuty metod redukcjonistycznych, które ujawniają się w przypadku badania zjawisk społeczno-gospodarczych będących z natury rzeczy skomplikowanymi i rekurencyjnymi (Lange, 1986, s. 96). Ten prosty argument o złożoności zjawisk społeczno-gospodarczych jest trudny do odrzucenia, szczególnie biorąc pod uwagę ilość indywidualnych

podmiotów biorących udział w realnych procesach oraz różnorodność warunków, w jakich procesy te zachodzą. Zgłębiając ten temat wskazać należy, że pomioty działają w warunkach *niepewności*, w rozumieniu Franka Knighta (1921), oraz *niepewności behawioralnej* i oportunistu, w rozumieniu Oliviera Williamsona (2000). W efekcie nie sposób skalkulować *a priori* ostatecznego wyniku działania faktycznej grupy podmiotów gospodarujących, bez wcześniejszej redukcji zmiennych obciążonych nieprzewidywalną *niepewnością*.

Ponieważ nie istnieje doskonała metoda badawcza, ani tym bardziej uniwersalna metoda badawcza nauk ekonomicznych, wszelkie metody redukcjonistyczne narażone są na zarzuty o zbytym uproszczeniu rzeczywistości lub pominięciu czynników istotnych, albo niedostrzeżeniu znaczenia emergentności (Anderson, 1972). Jednak mimo świadomości tego faktu ich użycie jest niezbędne do tworzenia ram teoretycznych i wyjaśniania wieloelementowych zjawisk. Z zakresu metodologii ekonomii znamienne jest stanowisko Friedmana (1953, s. 15), które można sprawdzić do tego, że nie tyle istotne są realistyczne założenia teorii, co jej zdolności do wyjaśniania realnych zjawisk. Podobnie Edward Lipiński (1981, s. 279–280) pisał, przy okazji krytyki modelu człowieka gospodarującego, że chociaż jest to koncepcja wysoce uproszczona i nieprawdziwa, to tkwi w niej jakaś ogólna doza prawdy. A zatem jest w niej coś, co pozwala dostrzec badaczowi sedno, istotę zjawiska, a w tym przypadku tę część natury ludzkiej, która w realiach gospodarczych odgrywa nadrzędną rolę. Kennetha J. Arrow i Frank Hahn (1991) również brali pod uwagę fakt, że analizowana przez nich równowaga ogólna dotyczy wyidealizowanej i zdecentralizowanej gospodarki, co oznacza, że jest ona bytem kontrfaktycznym. Nie zmienia to jednak ich stanowiska w kwestii tego, że ma ona wciąż znaczące walory poznawcze.

Oczywiście stosowanie metod redukcjonistycznych zawsze niesie ze sobą zagrożenie pominięcia istotnego czynnika, który diametralnie zmienia ostateczny kształt badanego zjawiska. Niemniej metody te są wciąż powszechnie stosowane zarówno pośrednio w procesie dydaktycznym, gdy teorii redukcjonistycznych się po prostu uczy, jak i bezpośrednio podczas badania nowych zjawisk. Podobnie jak w naukach fizykalnych – modele są niezastąpionym łącznikiem między abstrakcyjną strukturą matematyczną a faktyczną rzeczywistością (Cartwright, 1984, s. 159), tak w ekonomii modele są źródłem praw opisujących zachowania uczestników życia gospodarczego.

Na przestrzeni lat badacze zajmujący się metodologią nauk społecznych opracowywali i/lub korzystali z szeregu metod redukcjonistycznych. Dla Webera (1922, 1984, 1994, 2002) był to *typ idealny*, dla Leszka Nowaka (1972, 1977, 1980) była to metoda idealizacji i stopniowej konkretyzacji, dla Mäkiego (1992) była to izolacja, a z kolei Oskar Lange (1986) pisał o abstrakcjach ekonomii politycznej i procesie izolacji myślowej. Często zapomina się również o powszechnie przywoływanej przez ekonomistów zasadzie *cete-*

ris paribus, którą stosował w swojej metodzie badawczej Alfred Marshall (1920, s. 366), również postulując izolację pewnych czynników poprzez uznanie pozostałych za stałe. Jest ona niczym innym jak formą metody redukcjonistycznej, której chyba najpowszechniejszym przykładem zastosowania jest prawo opadających krzywych popytu. Do kanonu ekonomii przeszło ono w interpretacji Paula Samuelsona (1948) jako prawo malejącej konsumpcji, w którym wykorzystanie zasady *ceteris paribus* jest silnie zaakcentowane dla ukazania istoty odwrotnej relacji między ceną danego dobra a ilością zapotrzebowania na to dobro.

Trudno w zasadzie określić, kiedy metody redukcjonistyczne zdominowały teorię ekonomii ortodoksyjnej, jednak z pewnością duży wpływ miała na to wspomniana wcześniej myśl Webera, który literalnie określił swoją metodę jako *typ idealny*. Weber (1922; 1949) wprowadził pojęcie typu idealnego jako formę narzędzia badawczego, czy też analitycznego, będącego w pełni abstrakcyjnym bytem. Zdefiniował on *typ idealny* jako „obraz myślowy” (niem. Gedankenbilder) (Weber, 1922, s. 194), co w anglosaskiej tradycji określa się jako „konstrukt pojęciowy” (por. Weber, 1949, s. 93). Jest to zatem byt abstrakcyjny – konstrukt myślowy, któremu Weber przypisuje konkretne cechy lub konkretnych cech go pozbawia. Weberowski system definicji określający typ idealny jest rozległy, w zależności od funkcji poznawczej jaką ten ma pełnić w procesie badawczym. I tak, nie posiada on *cechy realności* względem rzeczywistości, a posiada on *cechę deformacji* względem tej samej rzeczywistości. Cecha deformacji ujawnia się w procesie tworzenia typu idealnego, a polega ona na jednostronnym uwydatnieniu konkretnego – lub konkretnych – punktów widzenia (Weber, 1922, s. 190). W efekcie nie można w rzeczywistym świecie znaleźć realnego obiektu, który będzie odpowiadał typowi idealnemu, gdyż ten jest jedynie i aż deformacją obiektów rzeczywistych. Zatem ani historyczna, ani terażniejsza rzeczywistość nie może przyjąć postaci tego rodzaju „obrazu myślowego” – czy może raczej „konstruktu myślowego”, gdyż nic nie stoi na przeszkodzie sformułowania go w oparciu o nieistniejące zjawiska.

Weber (1922, s. 194) dosłownie wskazuje, że to wyobrażenia badacza odpowiada za adekwatność punktów widzenia, które dokonują deformacji rzeczywistości i prowadzą do powstania typu idealnego. Wydawać by się mogło, że takie podejście nie ma natury naukowej, gdyż funkcja poznawcza jest zawarta w subiektywnej naturze ludzkiej intuicji. Nic jednak bardziej mylnego, gdyż u Webera adekwatność osądu badacza, co do punktów widzenia uwydatniających jedne cechy rzeczywistości względem innych, jest podyktowana obserwacją rzeczywistości, która ową adekwatność kształtuje. Podejście to jest unikalnym redukcjonizmem, gdyż faktycznie ogranicza ilość czynników, poprzez jednostronny punkt widzenia, ale nie polega ono na niwelacji w celu odkrycia istoty zjawiska, lecz utworzeniu linii demarkacyjnej, względem

której rzeczywistość można mierzyć, a zatem badać na ile rzeczywiste zjawiska odpowiadają typowi idealnemu.

W takim ujęciu typ idealny jest narzędziem poznania, a nie poznaniem samym w sobie. I to go w dużej mierze różni go od redukcjonizmu opartego na procesie eliminacji czynników pobocznych wpływających na badane zjawisko. Zauważyć należy, że w podejściu metodologicznym Webera typ idealny może być konstruktem nie tylko kontrfaktycznym, ale również konstruktem subiektywnym, zależnym od samego badacza. O ile fakt, że postawa badacza zmienia przedmiot badania (Maslow, 1990) jest powszechnie akceptowany, to Weber nie poszukuje tu mechanizmu uniwersalnego formułowania zasad tworzenia typu idealnego. Inaczej niż to czyni na przykład Nowak (1977), który wyraźnie poszukuje podstaw, według których rzeczywistość jest redukowana do poziomu pozwalającego dostrzec sedno zjawiska. Próba Nowaka polegająca na strukturyzacji procesu tworzenia modelu prowadzi do formalizacji procesu poznawczego, co sprzyja ustalaniu się pewnego rodzaju tradycji badawczej. Istotne z punktu widzenia niniejszych rozważań jest to, że badacze nawet skupieni w ramach tej samej szkoły nie czynią refleksji nad źródłem przyjętych schematów i sensem przyjmowanych metod – przyjmują je mimowolnie, by nie powiedzieć bezmyślnie.

Nie chodzi tu bynajmniej o kwestie poddawania osądowi ich słuszności, lecz o fakt, że badacze mogą nie zdawać sobie sprawy ze źródeł mechanizmów badawczych, których używają. O ile tworzenie modeli w ekonomii wydaje się naturalną formą uprawiania nauki instytucjonalnej, szczególnie w ortodoksji ekonomicznej, to umyka fakt, że jest to zakres metod redukcjonistycznych izolujących dane czynniki. Co więcej umyka w procesie rozwoju historycznego fakt, że Weber stosował typ idealny jako narzędzie poznania poprzez zestawianie zdeformowanego bytu z bytem realnym, a deformacji dokonywał w sposób *a priori*.

Fakt wpływu idei izolacji, czy też redukcjonizmu, na ekonomistów skupionych w ramach ortodoksji ekonomicznej uwidacznia się w ich działaniu i próby obrony pewnych założeń, których *de facto* bronić nie trzeba, gdyż są one ustalone na zasadzie wyboru metodologicznego. A zatem kontrfaktyczność modeli ekonomii ortodoksyjnej wynika z deformacji, która jest właśnie wyborem metodologicznym. A zatem pełna racjonalność podmiotów, czy też abstrahowanie od szeregu realnych czynników, nie wynika z braku zdolności poznawczych, lecz jest efektem wprowadzenia *typu idealnego*.

Dyskurs o miejscu nauki instytucjonalnej i rewolucji naukowych w świecie nauki z pewnością nie jest rozstrzygnięty, a poczynione wyżej spostrzeżenia mają bez wątpienia postać przyczynkarską. W każdym razie ideał nauki poszukującej prawdy nie oznacza braku nauki „pozornej” związanej ze zdobywaniem stopni i tytułów naukowych. Można zatem przyjąć, że celem badacza w nauce instytucjonalnej będzie kariera naukowa determinowana regulacjami awansowymi, w odróżnieniu od rewolucji naukowych. Nie ozna-

cza to, że nie poszukuje on prawdy – ten cel może mieć dla niego charakter drugorzędny, albo może być nawet niedostrzegalny. Sytuacja uległaby jeszcze większej komplikacji, gdyby uwzględnić kryteria określenia tego, czym właściwie jest prawda. Przyjmując spostrzeżenie McCloskey (1983), że celem ekonomistów nie powinno być poszukiwanie prawdy w ekonomii, a prawdy o ekonomii, można dojść do ujęć skrajnie relatywistycznych, w których badacz poszukuje zgodności ze swoim systemem postrzegania rzeczywistości. Idąc tutaj za wspomnianym wcześniej Friedmanem (1953, s. 15), badacz nauk ekonomicznych będzie wyjaśniał realne zjawiska, jednak jego system może mieć cechę deformacji względem rzeczywistości, czym z kolei zbliża się do weberowskiego *typu idealnego*. Co prawda nie jest to obszarem rozważań w niniejszym artykule, jednak ukazuje złożoność problematyki postrzegania ideału nauki w konfrontacji z pozornym celem nauki związanym z kształtowaniem kariery zawodowej.

3. UJĘCIE BEHAWIORALNE AKTYWNOŚCI BADAWCZEJ

O ile aspekty instytucjonalne są ogólnie przyjęte jako warunki naturalne środowiska akademickiego, które wpływają na kierunki rozwoju wiedzy, o tyle aspekt psychologiczny jest traktowany pobocznie. Chociaż wskazuje się na aspekty indywidualistyczne, dotyczące samego badacza jako źródło jego aktywności akademickiej, to jednak aspekty instytucjonalne, a głównie stosunków społecznych, dominują w badaniach. Oczywiście wskazuje się na rolę badacza (Znaniński, 1984) w procesie kształtowania wiedzy, jak również na jego subiektywny osąd wobec rzeczywistości (Maslow, 1990), ale kwestie poznania ogólnie przesuwa się na płaszczyznę epistemologii jako takiej, a samego badacza ponownie traktuje się jako byt abstrakcyjny.

Abraham Maslow (1990) we wstępie do swojej teorii motywacji poświęcił szczególną uwagę badaczowi jako podmiotowi subiektywnemu o indywidualistycznych atrybutach wpływających na jego podejście do przedmiotu badań. Jednak jego analiza nie była środkiem do zrozumienia samego badacza, lecz ukazaniem wpływu subiektywności badacza na kształt jego badań.

Wracając do problemu wyboru przed jakim stoi badacz w warunkach zastanej tradycji badawczej, należy odseparować warunki instytucjonalne od behawioralnych – tym samym traktując naukę instytucjonalną jako pewną stałą. W takim ujęciu proces socjalizacji podmiotu decyzyjnego w środowisku akademickim jest zakończony i zdeterminowany. Nie oznacza to jednak, że każdy badacz będzie podejmował takie same decyzje, gdyż na te będą wpływały zarówno osobnicze uwarunkowania, tradycja badawcza, jak i konkretna sytuacja, w warunkach której decyzja jest podejmowana.

W takim ujęciu badacza traktuje się indywidualistycznie, ale nie poza kontekstem instytucjonalnym. Czyli punktem centralnym nie jest przedmiot

badania (jakie teorie są rozwijane), ale podmiot badania i jego skłonności do podejmowania takich, a nie innych decyzji w stosunku do wyboru problemu badawczego i metody jego rozwiązywania. Można przyjąć, że wyborów tych można dokonać w dwojaki sposób – albo za pośrednictwem utartych schematów myślowych (przyjętych problemów badawczych i metod rozwiązywania łamigłówek intelektualnych), albo w sposób nieszablony (wychodząc poza schematy). Ujęcie to odpowiada współczesnemu podziałowi z zakresu nauk behawioralnych (Stanovich, West, 2000), które szeroko stosowane jest w ekonomii behawioralnej za sprawą głównie Daniela Kahnemana (2012) i jego interpretacji podziału mechanizmu myślenia człowieka na system 1 i system 2, co odpowiada kolejno myśleniu autonomicznemu i myśleniu refleksyjnemu. O ile system 1 (myślenie autonomiczne) ma relatywnie jasne granice i jest głęboko zakorzeniony w ewolucji gatunku ludzkiego, to system 2 (myślenie refleksyjne) nie jest ograniczone wyłącznie biologicznym dziedzictwem i może rozwijać się niezależnie od procesu ewolucji biologicznej¹. Co prawda człowiek, a zatem *homo sapiens*, posiada genetycznie uwarunkowanie do myślenia refleksyjnego (a tym samym do poznania analitycznego i sceptycyzmu krytycznego, czyli właśnie wychodzenia poza schematy), ale ponieważ system ten wymaga dużego wysiłku, to ma on znaczący udział w tym co biolodzy określają mianem budżetu energetycznego organizmów żywych (Begon et al., 2012). Łączy się to ściśle z teorią optymalnego żerowania, której podstawy opracowali Robert MacArthur i Eric Pianka (1966). Ich rozważania, paralelne do zasad efektywności instrumentalnej i teorii rozwoju prowadzonych na płaszczyźnie ekonomicznej, wskazują na optymalizację energii niezbędnej do zaspokojenia deficytu kalorycznego zwierząt. W świetle ich badań zasadnym jest twierdzenie, że człowiek jest z natury skąpcem poznawczym (Nęcka et al., 2020) i o ile nie ma do tego konkretnego bodźca to wykorzystywać będzie system 1 do rozwiązywania problemów, jako że jest mniej energochłonny.

Wydawać by się mogło, że zasada ta nie będzie właściwa dla środowiska akademickiego, gdyż badacze powinni być niejako zdeterminowani do myślenia analitycznego, a zatem powinni wykorzystywać system 2. Należy jednak pamiętać, że zawodowy charakter badaczy oznacza również konieczność zachowania standardu pracy akademickiej, niejednokrotnie związanej z maksymą „publikuj albo gin” (ang. *publish or perish*) oraz prostym faktem, że są to również ludzie i jak inni zdeterminowani są w dużym stopniu ograniczeniami biologicznymi.

W konsekwencji problemy badawcze rozwiązywane są w analogii do systemu 1, czyli z pomocą metod badawczych, które można porównać z heurystykami wyborów, w rozumieniu ekonomii behawioralnej. Przy czym należy

¹ Należy mieć na uwadze, że abstrahuje się tu od kwestii ewolucji kulturowej i kwestii zdolności myślenia człowieka, jako właściwości kolektywnej (społecznej), czego przykładem są kolektywne myślowe Flecka (1986).

tu zaznaczyć, że heurystyki badawcze, o których tu mowa, nie są heurystykami w znaczeniu ściśle takim samym jak zaproponowali je Amos Tversky i Kahneman (1974). Trudno mówić, że metody badawcze są zbiorem prostych zasad, które nie wymagają dużego wysiłku, jednak z pewnością są one bardziej efektywne w kształtowaniu kariery akademickiej. I o ile nie ulega wątpliwości, że stawianie problemów i ich rozwiązywanie nawet ogólnie przyjętą metodą nie jest tak automatyczne jak heurystyki identyfikowane w eksperymentach z zakresu analityki błędów poznawczych, to jednak na gruncie uprawiania nauki z pewnością wymagają one mniejszej dozy wysiłku badacza. Sam fakt, że stosuje się już znaną metodę, której etapy są jasno określone, pozwala uniknąć wysiłku związanego z przygotowaniem nowej procedury, która może okazać się błędna lub być odrzucona na poziomie recenzji pracy akademickiej. W konsekwencji wymagać może zdecydowanie więcej wysiłku, przy niepewnym sukcesie publikacyjnym. Wyrażając się prozaicznie - prościej jest badania replikować, niż opracowywać je od nowa. Nie oznacza to bynajmniej, że badacz stosujący znaną sobie metodę badawczą czy replikujący badania korzysta wyłącznie z systemu 1 i pozbawiony jest zupełnie refleksji nad badaniem, które wykonuje. Dzięki ulokowaniu w określonej tradycji badawczej i swojemu doświadczeniu naukowemu, pewne elementy procesu badawczego może poniekąd realizować automatycznie – chciałoby się rzec bez wysiłkowo.

To, co odróżnia heurystyki badawcze od heurystyk ogółem jest to, że te pierwsze są opracowane wcześniej w wysiłku intelektualnym, który raczej nazwać należy wysiłkowym, niż autonomicznym na zasadzie odruchu związanym z poznawczymi zdolnościami człowieka, wynikającymi z jego genów. Heurystyki w rozumieniu ogólnym mogą natomiast wynikać właśnie z samej własności aparatu poznawczego człowieka, lecz są one również weryfikowane doświadczalnie i osobniczo, na drodze uczenia się chociażby metodą prób i błędów lub obserwacji innych. W takim ujęciu analogia heurystyk badawczych do heurystyk ogólnych jest zasadna. Wypracowywane wcześniej są przyjmowane przez adeptów nauki w ich pracy akademickiej. Jeżeli przynoszą one efekty w procesie awansu badacza, to można uznać, że utrwalają się one jako pewne „uproszczone” zasady wnioskowania nad problemami badawczymi. Ich efektywność będzie się wyrażała prozaicznie, w postaci przyjętych publikacji lub uzyskiwanych stopni i tytułów naukowych. O ile stanowi to wielką zaletę heurystyk badawczych, którą zwykle kwituje się wolnym cytatem z Isaaca Newtona „jeśli widzę dalej, to dlatego, że stoję na ramionach gigantów”, to rodzi to też pewne obiektywne zagrożenia natury behawioralnej i instytucjonalnej.

W ujęciu behawioralnym prowadzić to może to naturalnego dla człowieka gospodarowania budżetem energetycznym – czyli uleganiu pokusie stosowania heurystyk, w tym badawczych, o ile nie ma konkretnych powodów do wykorzystania możliwości systemu 2, a więc głębszej refleksji nad danym

zagadnieniem. W ujęciu instytucjonalnym znaczenie mają z kolei czynniki społeczne legitymizujące istotność pewnych utartych metod rozwiązywania konkretnych problemów analitycznych (Fleck, 1986; Kuhn, 1968), co *de facto* przekłada się na ich użyteczność w procesie kształtowania kariery akademickiej.

Przykładem oddziaływania czynników behawioralnych, choć jedynie incydentalnym, jest anegdota przywołana przez Kahnemana (2012), który „wypróbował” wraz Tverskym ich formularz eksperymentu związany z przewidywaniem na podstawie reprezentatywności. Badacze stworzyli formularz opisujący fikcyjnego studenta (Tomasza W.) i określili zbiór możliwych kierunków studiów jakie mógł podjąć. W poleceniu prosili o ich uszeregowanie względem zgodności z opisem ów studenta. Formularz ten pokazali Robynowi Dawesowi zajmującemu się psychologią matematyczną, uznanemu badaczowi z zakresu badań na osądami i decyzjami. Dawes udzielił odpowiedzi opartej na stereotypie, typując Tomasza W. jako studenta informatyki, gdy w rzeczywistości bardziej prawdopodobne był, że studiuje on nauki humanistyczne lub społeczne. Wynika to z prostego faktu, że kierunki te są liczniejsze (a na pewno były w chwili przeprowadzania tego eksperymentu) i naukowiec, a szczególnie statystyk powinien wiedzieć, że jest to kwestia wartości bazowej i tym samym podejść do problemu analitycznie (wykorzystać system 2), a nie odpowiedzieć intuicyjnie (za pomocą heurystyk systemu 1 – w tym przypadku posługując się stereotypem studenta informatyki).

Powyższy przykład ukazuje, że nawet badacze podlegają „pokusie” wykorzystania systemu 1, co w pracy nad rozwiązywaniem problemów badawczych przyjmie postać *heurystyk badawczych*. Jeżeli skonfrontować to z dokonaniem konstruktivistów metodologicznych i wskazać na znaczenie społecznych uwarunkowań tworzenia nauki (por. Chalmers, 1997; Fleck, 1986; Kuhn, 1962; Motycka, 1980), to badacz będzie znajdował się pod presją korzystania z utartych schematów rozwiązywania problemów, co wspierane jest biologicznym imperatywem wynikającym z budżetu energetycznego. Stephen Gould (1991) komentując inny eksperyment Kahnemana i Tversky’ego (1983), który w istocie sprowadzał się do tego samego problemu stereotypu, stwierdził, że mimo wiedzy na jego temat nie mógł się oprzeć silnej pokusie skorzystania z rozwiązania heurystycznego i tym samym nie zadawać sobie trudu myślenia refleksyjnego.

Ujęcie behawioralne, choć formalnie odseparowane od ujęcia instytucjonalnego, jest wobec niego komplementarne. System metodycznego kształtowania sposobów rozwiązywania problemów badawczych od systemu kształcenia do ulokowania w określonej tradycji badawczej kreuje mechanizmy korzystania z *heurystyk badawczych*. Zatem wraz z postępem wiedzy badacza, jego doświadczenia naukowego i znajomości problemów i metod badawczych, wzrasta możliwość szerokiego wykorzystywania tych *heurystyk*

badawczych w uprawianiu nauki. Nie oznacza to bynajmniej, że badacze dążą do pomijania myślenia analitycznego w ramach rozwiązywania problemów badawczych. Jak już zostało zauważone, z idealistycznej perspektywy nauka wiąże się z poszukiwaniem prawdy. Natomiast z tej pragmatycznej badacz funkcjonuje w określonych warunkach instytucjonalnych, które wymagają od niego spełniania specyficznych wymogów mogących streścić się w maksymie „publikuj albo zgiń”. Wynika z tego, że zdobywanie stopni i tytułów naukowych jest w pewnym zakresie pochodną uprawiania nauki i to niezależnie od tego, czy stosuje w badaniach niemal wyłącznie *heurystyki badawcze*, czy też podejmuje się głębszej refleksji nad obranym problemem badawczym.

4. PODSUMOWANIE

W artykule podjęto próbę odpowiedzi na pytanie: dlaczego ekonomiści stosują metodę izolacji do rozwiązywania problemów ekonomicznych? Jednocześnie punktem wyjścia była hipoteza, zgodnie z którą naukowcy posługują się *w badaniach naukowych heurystykami w ramach nauki instytucjonalnej*. Przyjęcie w ten sposób sformułowanej hipotezy badawczej oznacza, że wybory metodologiczne są uwarunkowane zarówno czynnikami instytucjonalnymi jak i behawioralnymi.

Nauka instytucjonalna opiera się na utrwalonych paradygmatach, czego praktycznym wyrazem są konkretne szkoły i tradycje badawcze, które ograniczają zakres możliwych do postawienia tez, tym samym samoczynnie chroniąc zastany porządek. W tym kontekście metoda izolacji jako forma redukcjonizmu, jest nie tylko narzędziem poznawczym, ale również bezpieczną i akceptowalną strategią w ramach akademickiego *status quo*. Wpisuje się ona w tradycję badawczą nauk ekonomicznych, a przez to zapewnia replikacyjne realizowanie badań, którego naturalną konsekwencją jest zdobywanie stopni i tytułów naukowych. By jednak nie trywializować izolacji jako narzędzia i środka do celu, należy podkreślić, że jej powszechność w ekonomii ortodoksyjnej wynika z potrzeby uproszczenia złożonych zjawisk społeczno-gospodarczych. I z tej perspektywy trudno odmówić jej walorów poznawczych.

W ujęciu behawioralnym wskazano natomiast, że *badacz*, niezależnie jak bardzo by się chciało nadać mu górnolotnego znaczenia, jest poznawczo ograniczony, co sprzyja stosowaniu heurystyk, czyli uproszczonych reguł decyzyjnych. W kontekście pracy badawczej i biorąc pod uwagę funkcjonalne warunki środowiska akademickiego, wybór znanej i sprawdzonej metody jaką jest izolacja, może być traktowany jako forma *heurystyki badawczej*, zatem efektywnej, choć niekoniecznie optymalnej z punktu widzenia poznania naukowego, metody rozwiązywania problemów badawczych. Jak to zo-

stało zauważone, analogicznie do systemu 1 w teorii Kahnemana, badacze mogą upraszczać swoje wybory badawcze (problem badawczy) i metodologiczne (metoda badawcza), szczególnie w warunkach presji instytucjonalnej sprowadzającej się do maksymy „publikuj albo gin”, której naturalnym symptomem jest zdobywanie stopni i tytułów naukowych.

W każdym razie dominacja metody izolacji w ekonomii nie wynika wyłącznie z jej walorów poznawczych, lecz w dużej mierze z uwarunkowań instytucjonalnych (nauka instytucjonalna) i behawioralnych (*heurystyki badawcze*). Wybory metodologiczne adeptów nauk ekonomicznych są zatem efektem interakcji między tradycją badawczą, ograniczeniami poznawczymi a imperatywami kariery akademickiej. W tym sensie hipoteza o stosowaniu heurystyk w nauce instytucjonalnej znajduje swoje uzasadnienie. Ekonomisci jako podmioty behawioralne funkcjonujące w określonym środowisku akademickim, skłonni są do stosowania uproszczonych metod w swojej pracy akademickiej. Tym samym zastana metoda izolacji nie jest poddawana ciągłej krytyce i mogła się utrwalić instytucjonalnie. Nie oznacza to braku postępu ani braku podejmowania wysiłku w celu przewyciężenia ograniczeń instytucjonalnych, ale wyjaśnia fakt stosowania utartych metod w pracy naukowej. Należy również podkreślić, że w niniejszym artykule, nie przyjęto za cel wykazania zdegenerowania się metod opartych na izolacji, nawet biorąc pod uwagę osiągnięcia z zakresu informatyki i możliwości operowania na dużych zbiorach zmiennych istotnych. Wynika to z tego, że metody te wciąż pozostają aktualne i posiadają walory poznawcze. Podobnie jak heurystyki decyzyjne w ekonomii behawioralnej, tak *heurystyki badawcze* zwykle trafnie rozwiązują problemy badawcze, co sprzyja karierze akademickiej. Niemniej mogą powodować analogicznie do heurystyk decyzyjnych błędy poznawcze, które w tym przypadku przyjmują formę trwania przy błędnych teoriach.

Nie ulega także wątpliwości, iż refleksja nad sposobami uprawiania nauki jest nieodzownym elementem pracy badacza i dzieje się to niezależnie od tego, czy mówimy o nauce instytucjonalnej czy rewolucji naukowej. Zatem zarówno realizowanie wyidealizowanego celu nauki w postaci poszukiwania prawdy ma znaczenie w pracy badawczej, jak i wypełnianie pozornego jej celu, którym jest spełnianie się w karierze akademickiej (zdobywanie stopni i tytułów naukowych).

BIBLIOGRAFIA

- G.A. Akerlof, *Writing the “The Market for ‘Lemons’”: A Personal Interpretive Essay*, NobelPrize.Org, 2003, <https://www.nobelprize.org/prizes/economic-sciences/2001/akerlof/article/>
- P.W. Anderson, *More Is Different: Broken symmetry and the nature of the hierarchical structure of science*, *Science*, 177(4047), 1972, 393–396. <https://doi.org/10.1126/science.177.4047.393>

- K.J. Arrow, F.H. Hahn, *General competitive analysis*. North-Holland, Amsterdam/New York/Oxford 1991.
- Arystoteles, *Topiki. O dowodach sofistycznych*, Państwowe Wydawnictwo Naukowe, Warszawa 1978.
- M. Begon, C.R. Townsend, J.L. Harper, *Ecology: From individuals to ecosystems*, Blackwell, Malden 2021.
- P.L. Berger, *Invitation to sociology: A humanistic perspective*, Anchor Books, New York 1963.
- M. Blaug, *Metodologia ekonomii*, Wydawnictwo Naukowe PWN, Warszawa 1995.
- G. Burrell, G. Morgan, *Sociological Paradigms and organisational analysis: Elements of the sociology of corporate life*, Heinemann Educational Books, London 1979.
- A. Calaprice, S. Elworthy, T. Lipscombe, R. Penrose, J.J. Stachel, *5 prac, które zmieniły oblicze fizyki*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa 2008.
- D.T. Campbell, *Assessing the impact of planned social change*, Evaluation and Program Planning, 2(1), 1979, 67–90. [https://doi.org/10.1016/0149-7189\(79\)90048-X](https://doi.org/10.1016/0149-7189(79)90048-X)
- Cartwright, N. (1984). How the laws of physics lie (Repr). Clarendon Press.
- A.F. Chalmers, *Czym jest to, co zwiemy nauką?: Rozważania o naturze, statusie i metodach nauki; wprowadzenie do współczesnej filozofii nauki*, Siedmioróg Wrocław 1997.
- L. Fleck, *Zagadnienie teorii poznawania*, Przegląd filozoficzny, Rocznik XXXIX, 1936, 3–37.
- _____, *Powstanie i rozwój faktu naukowego: Wprowadzenie do nauki o stylu myślowym i kolektywnie myślowym*, Wydawnictwo Lubelskie, Lublin 1986.
- M. Friedman, *Essays in positive economics*, University of Chicago Press, Chicago 1953.
- S.J. Gould, *Bully for Brontosaurus: Reflections in Natural History*. W. W. Norton & Company, New York/London 1991.
- A. Grobler, *Prawda a względność*, Aureus, Kraków 2000.
- L. Hardt, *Studia z realistycznej filozofii ekonomii*, Wydawnictwo C.H. Beck, Warszawa 2013.
- D. Kahneman, *Pułapki myślenia: O myśleniu szybkim i wolnym*, Media Rodzina, Poznań 2012.
- I. Kant, *Krytyka praktycznego rozumu*, Wydawnictwo Naukowe PWN, Warszawa 2012.
- F. Knight, (1921). Riskuncertainty.pdf. The Riverside Press.
- T.S. Kuhn, *The structure of scientific revolutions*, University of Chicago Press, Chicago 1962.
- _____, *Struktura rewolucji naukowych*, Państwowe Wydawnictwo Naukowe, Warszawa 1968.
- _____, *The essential tension: Selected studies in scientific tradition and change*, University of Chicago Press, Chicago 1977.
- _____, *Struktura rewolucji naukowych*, Wydawnictwo Aletheia, Warszawa 2009.
- I. Lakatos, *The Methodology of Scientific Research Programmes*, in: J. Worrall, G. Currie (eds.), *Philosophical Papers*, Cambridge University Press, Cambridge 1978. <https://doi.org/10.1017/CBO9780511621123>
- O. Lange, *Ekonomia polityczna*, Państwowe Wydawnictwo Naukowe, Warszawa 1986.
- H. Leibenstein, *Poza schematem homo oeconomicus: Nowe podstawy mikroekonomii*, Państwowe Wydawnictwo Naukowe, Warszawa 1988.
- E. Lipiński, *Problemy, pytania, wątpliwości: Z warsztatu ekonomisty*, Państwowe Wydawnictwo Ekonomiczne, Warszawa 1981.
- R.H. MacArthur, R.E. Pianka, *On Optimal Use of a Patchy Environment*, The American Naturalist, 100(916), 1966, 603–609. <https://doi.org/10.1086/282454>
- U. Mäki, *On the method of isolation in economics*, Poznan Studies in the Philosophy of the Sciences and the Humanities, 26, 1992, 319–354.
- A. Marshall, *Principles of Economics*, Macmillan and Co., London 1920.
- A.H. Maslow, *Motywacja i osobowość*, Instytut Wydawniczy Pax, Warszawa 1990.
- D.N. McCloskey, *The Rhetoric of Economics*, Journal of Economic, 21(2), 1983, 481–517.
- A. Motyka, *Relatywistyczna wizja nauki. Analiza krytyczna koncepcji T. S. Kuhna i S. E. Toulmina*, Zakład Narodowy Imienia Ossolińskich Wydawnictwo Polskiej Akademii Nauk, Wrocław/Warszawa/Kraków/Gdańsk 1980.
- E. Nęcka, S. Wichary, J. Orzechowski, B. Szymura (Red.), *Psychologia poznawcza*, PWN, Warszawa 2020.
- L. Nowak, *Model ekonomiczny. Studium z metodologii ekonomii politycznej*, Państwowe Wydawnictwo Ekonomiczne, Warszawa 1972.
- _____, *Wstęp do idealizacyjnej teorii nauki*, Państwowe Wydawnictwo Naukowe, Warszawa 1977.
- _____, *The Structure of Idealization*, Springer Netherlands. Dordrecht 1980. <https://doi.org/10.1007/978-94-015-7651-2>
- S. Pinker, *Racjonalność: Co to jest, dlaczego jej brakuje, dlaczego ma znaczenie*, Zysk i s-ka Wydawnictwo, Poznań 2021.
- K. Popper, *The logic of scientific discovery*, Routledge, London 2010.

- E.S. Reich, *Plastic fantastic: How the biggest fraud in physics shook the scientific world*, Palgrave Macmillan, Basingstoke (Hampshire) 2009.
- G. Ritzer, *Sociology: A Multiple Paradigm Science*, *American Journal of Sociology*, 81(6), 1975, 156–167. <https://doi.org/10.1086/226243>
- W. Sady, *Spór o racjonalność naukową. Od Poincarégo do Laudana*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń 2013.
- P.A. Samuelson, *Economics. An Introductory Analysis*, McGraw-Hill Book Company., Inc., New York/Toronto/London 1948.
- R.F. Service, *Bell Labs Fires Star Physicist Found Guilty of Forging Data*, *Science*, 298(5591), 2002, 30–31. <https://doi.org/10.1126/science.298.5591.30>
- K.E. Stanovich, R.F. West, *Individual differences in reasoning: Implications for the rationality debate?*, *Behavioral and Brain Sciences*, 23(5), 2000, 645–665. <https://doi.org/10.1017/S0140525X00003435>
- N. Taleb, *Czarny łabędź: Jak nieprzewidywalne zdarzenia rządzą naszym życiem*, Zys i S-ka Wydawnictwo, Poznań 2020.
- A. Tversky, D. Kahneman, *Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty*, *Science*, 185(4157), 1974, 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- _____, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, *Psychological Review*, 90(4), 1983, 293–315. <https://doi.org/10.1037/0033-295X.90.4.293>
- M. Weber, *Gesammelte Aufsätze zur Wissenschaftslehre*. Verlag von J. C. B. Mohr (Paul Siebeck), Tübingen 1922.
- _____, *The Methodology of the Social Sciences*, The Free Press, Glencoe 1949.
- _____, *Szkice z socjologii religii*, Książka i Wiedza, Warszawa 1984.
- _____, *Etyka protestancka a duch kapitalizmu*, Wydawnictwo „Test”, Lublin 1994.
- _____, *Gospodarka i społeczeństwo. Zarys socjologii rozumiejącej*, Wydawnictwo Naukowe PWN Warszawa 2002.
- O.E. Williamson, *The New Institutional Economics: Taking Stock, Looking Ahead*, *Journal of Economic Literature*, 38(3), 2000, 595–613. <https://doi.org/10.1257/jel.38.3.595>
- F. Znaniecki, *Spoleczne role uczonych*, Państwowe Wydawnictwo Naukowe, Warszawa 1984.

ON RESEARCH HEURISTICS IN INSTITUTIONAL SCIENCE FROM THE PERSPECTIVE OF ECONOMIC SCIENCES

ABSTRACT

The article addresses the issue of the dominance of the isolation method in economic sciences within the context of institutional and behavioral conditions. The authors hypothesize that researchers employ research heuristics within institutional science, understood in Kuhn's terms as a period of paradigmatic consensus. The isolation method, as a form of reductionism derived from Weber's concept of the ideal type, serves not only as a cognitive tool simplifying complex socio-economic phenomena but also as a safe strategy for shaping one's academic career. From a behavioral perspective, the researcher, as a cognitively limited agent, tends to use proven methods that can be compared to decision-making heuristics known from behavioral economics. Institutional pressure expressed in the maxim "publish or perish" and the biological imperative of managing the energy budget foster the use of established research patterns. The article demonstrates that economists' methodological choices result from the interaction between research tradition and cognitive constraints, as well as the requirements of an academic career.

Keywords: institutional science, isolation method, research heuristics, methodology of economics.

O AUTORACH:

Sławomir Czetwertyński — dr hab., Katedra Mikroekonomii i Ekonomii Instytucjonalnej, Uniwersytet Ekonomiczny we Wrocławiu, ul. Komandorska 118/120, Wrocław.

Email: slawomir.czetwertynski@ue.wroc.pl

Jakub Marcinkowski — dr hab., Katedra Zarządzania Strategicznego i Logistyki, Uniwersytet Ekonomiczny we Wrocławiu, ul. Komandorska 118/120, Wrocław.

Email: jakub.marcinkowski@ue.wroc.pl

INFORMACJE DLA AUTORÓW

Przygotowanie tekstów

1. Przyjmujemy teksty rozpraw i studiów do 1,5 arkusza wydawniczego (60 000 znaków ze spacjami), polemik i głosów w dyskusjach – do 0,5 arkusza (20 000 znaków ze spacjami), recenzji – do 0,4 arkusza (około 16 000 znaków ze spacjami). W uzasadnionych przypadkach dopuszczamy wyjątki. Należy je uzgodnić wcześniej z zespołem redakcyjnym.

2. Prosimy autorów o przysyłanie tekstów w edytorze Word 1997–2003, z przypisami dolnymi, a nie końcowymi.

2a. Do każdego tekstu powinno zostać dołączone streszczenie w jęz. polskim (zamieszczone na początku tekstu) oraz w jęz. angielskim (na końcu tekstu), oraz słowa kluczowe w jęz. angielskim, informacja o afiliacji autora (umieszczona pod imieniem i nazwiskiem autora).

2b. Pożądane jest dzielenie tekstu na zatytułowane rozdziały.

3. Cytowanie pozycji literatury powinno zostać przygotowane według poniższego schematu: **Monografie:** Max Scheler, *Problemy socjologii wiedzy*, przeł. Stanisław Czerniak *et al.*, PWN, Warszawa 1990, s. 32.

Artykuły w czasopismach: Nelson Goodman, What Should Not Be Said about Representation?, *Journal of Aesthetics and Art Criticism*, 1987–1978, 46, s. 419–425.

Rozprawy w monografiach zbiorowych: E. Mayr, Die Darwinsche Revolution und die Widerstände gegen die Selektionstheorie, w: J. Herbig, R. Hohlfeld (red.), *Die zweite Schöpfung. Geist und Ungeist in der Biologie des 20. Jahrhunderts*, Hanser, München 1990, s. 44–70.

Odsyłacze do literatury należy umieszczać na jeden ze dwóch sposobów:

A) w przypisach dolnych;

B) w zamieszczonej na końcu tekstu **Bibliografii**. W takim przypadku odsyłacze do literatury powinny być umieszczone w tekście według następującego schematu: nazwisko autora, rok wydania, strony, na przykład: (Giere, 1988, s. 25).

Wybrany przez Autora sposób A) lub B) powinien być stosowany konsekwentnie w całym tekście.

C) Bibliografia winna być uporządkowana alfabetycznie, według nazwisk autorów.

4. Elementy tekstu, które Autor pragnie wyróżnić, należy pisać rozstrzelonym drukiem.

5. Tytuły i podtytuły – wypośrodkowane, półgrubą czcionką.

6. Notki (przypisy) – dolne, a nie końcowe.

7. Autorzy proszeni są o przygotowanie tekstu do celów peer-blind review, czyli o niezamieszczenie w tekście informacji pozwalających zidentyfikować autora. Dane autora na pierwszej stronie tekstu zostaną usunięte przez redakcję przed przekazaniem jej recenzentom.

8. Autorzy są ponadto proszeni o ujawnienie wszystkich osób biorących udział w powstawaniu publikacji oraz ewentualnych źródeł powstawania publikacji. To rozwiązanie zastosowane przez redakcję ma zabezpieczać publikacje przed zjawiskiem ghost-writing.

9. Autorzy są też proszeni o złożenie deklaracji (także elektronicznie, w formie skanu z podpisem), że tekst przysyłany do druku nie jest przedrukiem tekstu wcześniej publikowanego.

10. Materiały należy przysłać pocztą elektroniczną na adres:

filozofia.nauka@ifispan.waw.pl

11. Ewentualne diagramy, ryciny i inne formy graficzne znajdujące się w tekstach powinny być czarno-białe.

12. Wzory matematyczne powinny być zapisane w formie Word. W razie trudności możliwe są indywidualne negocjacje z redakcją.

Proces recenzowania

Teksty nadsyłane do czasopisma są recenzowane zgodnie ze standardami peer-blind review. Szablon recenzji oraz lista recenzentów każdego wydanego tomu czasopisma jest podana na stronie internetowej czasopisma. Lista recenzentów nie jest stała. Redakcja powołuje recenzentów w zależności od tematyki przysyłanych tekstów. Daje to gwarancję oceniania tekstów przez faktycznych specjalistów problematyki rozważanej w nadsyłanych tekstach.

Zakres tematyczny

W czasopiśmie jest prezentowana cała filozoficzna problematyka, która ma związki z nauką, a więc

- problematyka filozoficzna asymilująca wyniki nauki jako przedmiot swych analiz, źródła informacji lub inspiracje;
- epistemologia i metodologia;
- dociekania nad filozofią projektowaną jako nauka;
- rozważania nad relacjami pomiędzy nauką a światem życia, rzeczywistością społeczną i kulturą.

Interdyscyplinarność

Publikujemy także prace łączące wątki *stricte* filozoficzne z typowo naukowymi. Zamierzenie to ujmuje rozmycie i płynność granic pomiędzy nauką i filozofią.

Multiprogramowość

Nie wprowadzamy programowych metafizycznych ograniczeń. Czasopismo nie jest forum jednej tylko szkoły filozoficznej. Multiprogramowość jest promowana w czasopiśmie między innymi jako wyraz specyfiki obecnej filozofii.