DOI: https://doi.org/10.18372/2412-2157.1.21227
УДК 165.6/7:001.8:316.77:004.9 (045)

ЕПІСТЕМІЧНА АРХІТЕКТУРА БЕЗПЕКИ: ФІЛОСОФСЬКО-МЕТОДОЛОГІЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ ІНФОРМАЦІЙНОЇ ТА КОГНІТИВНОЇ БЕЗПЕКИ В УМОВАХ КІБЕРСУСПІЛЬСТВА

Геннадій Христокін,[1]
Ірина Верховцева, [2]
Рена Марутян [3]

Державний університет «Київський авіаційний інститут»
https://orcid.org/0000-0002-2663-3055 [1]
https://orcid.org/0000-0002-5682-993X [2]
Київський національний університет імені Тараса Шевченка [3]
https://orcid.org/0000-0001-9184-1590

Анотація. *У статті обґрунтовується концепт епістемічної архітектури безпеки як розвиток і поглиблення ідеї структурної кореляції класичної, некласичної і постнекласичної раціональності. Під структурною кореляцією раціональностей розуміється функціональна взаємодоповнюваність класичної (орієнтація на факт), некласичної (увага до позиції спостерігача) та постнекласичної (системний аналіз умов знання) раціональностей. Чотириполюсна архітектура когнітивної та інформаційної безпеки має: верифікаційний, комунікаційний, наративний та ідентифікаційний виміри. Кожен з них має власну вразливість і власну функцію захисту. Чотири виміри взаємно утримують один одного, і руйнування будь-якого з них робить архітектуру безпеки нестабільною. Штучний інтелект аналізується як перший актор, що атакує всі чотири рівні одночасно і ефективно. Вводиться поняття епістемічної резилієнтності як здатності суспільства відновлювати цю архітектуру після атаки.*

Ключові слова: епістемічна архітектура безпеки, когнітивна безпека, інформаційна безпека, структурна кореляція раціональностей, епістемічна резилієнтність, наративна безпека, комунікаційний вимір, штучний інтелект, кіберсуспільство, деінструменталізація раціональності.

Вступ

Питання про те, як захистити суспільство від інформаційних атак, стало сьогодні одним із найактуальніших у безпековому дискурсі. Але за цим практичним питанням стоїть глибше і менш очевидне: а що саме ми захищаємо? Повномасштабна збройна агресія Росії проти України зробила цю проблему гранично конкретною. Кібератаки на критичну інфраструктуру, масовані кампанії дезінформації, наративна війна на міжнародній арені – всі ці явища є проявами єдиної стратегії, яку не можна описати лише в технічних або правових категоріях. За ними стоїть щось глибше: систематична атака на умови, за яких суспільство здатне себе відтворювати і захищати, на спроможність суспільства мислити, відрізняти реальне від сконструйованого, підтримувати спільні смислові рамки і зберігати волю до колективної дії. Саме виклики і наслідки цієї атаки є предметом нашого дослідження.

У попередній роботі ми обґрунтували концепт структурної кореляції класичної, некласичної і постнекласичної раціональності як критерій когнітивної безпеки (Христокін, Верховцева, Слюсар 2025). Ця стаття є продовженням і поглибленням тієї концепції. Тут ми рухаємось далі: від структурної кореляції раціональностей до поняття епістемічної архітектури безпеки, і від опису вразливості раціональностей до концепту епістемічної резилієнтності.

У літературі, присвяченій інформаційній і когнітивній безпеці, склалася певна асиметрія. Технічні та правові підходи розробляють інструменти захисту мереж, стандарти кіберстійкості, регуляторні рамки (Rid 2013). Безпекові студії описують загрози: дезінформацію, когнітивну війну, гібридні операції (Paul and Matthews 2016; Du Cluzel 2020). Але філософського пояснення того, чому ці загрози є системно ефективними, тобто пояснення архітектури вразливості – наразі бракує. Ані технічна, ані правова традиція не може відповісти на питання: чому суспільство, захищене технічно, може програти інформаційну війну?

Відповідь на це питання шукали в різних напрямах. Ю. Габермас показав, що легітимність демократичних рішень залежить від умов раціональної комунікації: «раціональна дискусія є не просто засобом досягнення консенсусу, а умовою, за якої консенсус може мати нормативну силу» (Habermas 1984, 17). М. Фуко описав режими істини як механізми, через які влада конституює знання (Foucault 1980). М. Кастельс зафіксував, що в мережевому суспільстві влада дедалі більше визначається здатністю контролювати інформаційні потоки (Castells 2009). Л. Флоріді запропонував концепт інфосфери як нового онтологічного середовища (Floridi 2014). Кожен із цих підходів описує частину реальності, але жоден не дає інтегрованої архітектурної рамки.

Окрема лінія досліджень пов'язана з наративним виміром інформаційної безпеки. Конфлікт гранд-наративів, боротьба за контроль над символічним порядком, деконструкція ідеологічних конструктів показують, що інформаційна війна є передусім боротьбою за те, яка версія реальності стане легітимною (Христокін 2026). Паралельно дослідження когнітивних упереджень (Kahneman and Tversky 1979) відкрили рівень ідентифікаційних механізмів, через які маніпуляція стає можливою.

Проте жоден із відомих нам підходів не пропонує інтегрованої філософської рамки, яка б одночасно описувала чотири виміри епістемічної вразливості суспільства, пояснювала їх необхідну взаємозалежність і давала операційний критерій епістемічної резилієнтності. Саме цю прогалину намагається заповнити дана стаття.

Мета та завдання дослідження

Мета статті – обґрунтувати концепт «епістемічної архітектури безпеки» як інтегрованої

філософсько-методологічної концепції для дослідження інформаційної і когнітивної безпеки в умовах кіберсуспільства. Це конкретизується в чотирьох взаємопов'язаних завданнях:

1) виявити та описати виміри епістемічної архітектури безпеки та обґрунтувати їхню взаємну необхідність?

2) описати специфіку загроз для кожного рівня і показати, чому руйнування будь-якого з них дестабілізує архітектуру безпеки в цілому;

3) обґрунтувати поняття епістемічної резилієнтності як здатності суспільства відновлювати цю архітектуру;

4) проаналізувати загрози з боку штучного інтелекту для рівнів епістемічної архітектури безпеки**.**

Методологія дослідження.

Методологічна позиція даної статті має своїм відправним пунктом типологію В. Стьопіна, яку ми використовуємо критично як евристику. Соціальна епістемологія Г. Лонгіно дозволяє концептуалізувати об'єктивність як інституційну практику, а не властивість індивідуального методу. Теорія комунікативної дії Ю. Габермаса постачає інструментарій для розрізнення легітимної аргументації і стратегічної маніпуляції (Habermas 1984). Центральним методологічним інструментом дослідження є аналіз залежності між чотирма вимірами епістемічної архітектури безпеки. Наративний аналіз і дискурс-аналіз задіяні для опису наративного виміру архітектури безпеки (Ricoeur 1990; Foucault 1980). Теорія когнітивних упереджень Д. Канемана і А. Тверськи (Kahneman and Tversky 1979), а також соціальний конструктивізм П. Бергера і Т. Лукмана (Berger and Luckmann 1966) постачають інструменти для аналізу ідентифікаційного виміру. Для аналізу комунікаційного виміру залучаються концепції медіаархітектури Т. Гіллеспі (Gillespie 2018) та дослідження мережевої пропаганди Й. Бенклера, Р. Фаріса і Х. Робертса (Benkler, Faris, and Roberts 2018), які показують, що структура інформаційного середовища є самостійним чинником епістемічної вразливості, незалежним від змісту конкретних повідомлень. Для аналізу загроз з боку штучного інтелекту використовуються підходи філософії ШІ і дослідження алгоритмічних середовищ (Floridi 2014; Russell 2019).

Результати дослідження

У попередній статті (Христокін, Верховцева, Слюсар 2025) ми показали, що класична, некласична і постнекласична раціональність є не послідовними етапами, а функціонально різними відповідями на різні пізнавальні проблеми. Між ними існує структурна кореляція: кожен тип виконує незамінну функцію перевірки двох інших, і руйнування будь-якого полюсу веде до специфічних деформацій знання. Ця ідея є відправним пунктом для нинішнього аналізу.

Але структурна кореляція раціональностей описує передусім рівень наукового і академічного дискурсу. Якщо ми хочемо говорити про когнітивну безпеку суспільства в цілому, нам потрібна візія, яка описує те, як ця структура функціонує в публічному просторі, в умовах масових комунікацій, алгоритмічних середовищ і цілеспрямованих інформаційних атак.

Епістемічна архітектура безпеки (epistemic security architecture) – це не набір заходів і не інституційна схема. Це опис необхідної внутрішньої структури суспільного пізнання, яка дозволяє суспільству відрізняти реальне від сконструйованого, утримувати спільні смислові рамки і зберігати здатність до колективної дії на основі надійного знання.

Ми пропонуємо розрізняти чотири виміри суспільного пізнання, кожен з яких відповідає на одне з чотирьох ключових питань: що є фактом? Яким чином ми про нього дізнаємося і через які канали? Як ми інтерпретуємо факти і яким ми бачимо світ через наративи і символи? Хто ми такі в світлі цих наративів? Перше питання є верифікаційним, друге є комунікаційним, третє – наративним, четверте постає ідентифікаційним. Так виникає чотириполюсна архітектура, що має верифікаційний, комунікаційний, наративний та ідентифікаційний виміри. Жодне з них не зводиться до іншого: відповідь на питання про факти не замінює питання про канали їх поширення, а інтерпретація не замінює ідентифікацію. Проаналізуємо їх по черзі.

Верифікаційний вимір відповідає за встановлення фактів. Це рівень, на якому суспільство питає: що насправді відбулося? Який доказ? Хто підтверджує? Цей вимір відповідає класичній раціональності – орієнтації на знання, незалежне від позиції спостерігача. Його інституційне вираження – це незалежна журналістика зі стандартами верифікації, факт-чекінг, наукова експертиза, судові процедури встановлення істини.

У цифровому середовищі цей вимір стає одночасно найважливішим і найбільш атакованим. Алгоритмічні системи оптимізовані не на точність, а на залучення, і саме тому класичний орієнтир «чи це відповідає дійсності?» поступається перед критерієм «чи це викликає емоційну реакцію?». С. Восуґі, Д. Рой і С. Арал показали, що фейкові новини поширюються в 6 разів швидше за правдиві (Vosoughi, Roy, and Aral 2018, 1147). Воєнні злочини в Бучі є верифікованим фактом, підтвердженим незалежними журналістами, супутниковими знімками і міжнародними слідчими. Але і цей факт стає об'єктом систематичного заперечення і саме тому його захист є питанням безпеки, а не лише академічної доброчесності.

Комунікаційний вимір відповідає за структуру середовища, через яке інформація рухається між людьми. Це рівень, на якому суспільство питає: яким чином ми про це дізнаємося? Через які канали, платформи, алгоритми? Хто контролює видимість повідомлень? Цей вимір є якісно відмінним від верифікаційного. Він не питає, чи є повідомлення правдивим, а питає, чи є воно взагалі чутним і в якому контексті.

Тарлтон Гіллеспі точно зафіксував, що платформи не є нейтральними провідниками контенту. Через алгоритми модерації і дизайн

інтерфейсу вони «визначають форму публічного дискурсу і вони це знають» (Gillespie 2018, 25). Це означає, що сама архітектура комунікаційного середовища є вже формою епістемної влади, незалежно від конкретних повідомлень, що циркулюють у ній. Й. Бенклер, Р. Фаріс і Х. Робертс у масштабному дослідженні мережевої пропаганди показали: асиметрії в поширенні дезінформації пояснюються не стільки змістом повідомлень, скільки мережевою архітектурою медіасередовища (Benkler, Faris, and Roberts 2018, 8). Певні мережі структурно сприяють поширенню неверифікованого контенту і це є інфраструктурною, а не лише змістовою проблемою.

Комунікаційний вимір є тим рівнем, де вразливість найменш помітна. Людина може мати доступ до нараправдивої інформації, але якщо архітектура платформи систематично знижує її видимість порівняно з маніпулятивним контентом, верифікаційний вимір залишається безсилим. Саме тому атака на комунікаційний вимір є особливо витонченою. Вона не фальсифікує факти, а керує потоками, крізь які факти досягають або не досягають аудиторії. Захист цього виміру вимагає прозорості алгоритмів, різноманітності медіа-платформ і регуляторних механізмів, спрямованих не проти конкретного контенту, а проти маніпулятивних архітектур розповсюдження.

Наративний вимір відповідає за інтерпретацію фактів. Це рівень, на якому суспільство питає: що це означає? Яка версія подій є легітимною? Хто є «своїм» і хто є «чужим»? Наратив є не прикрасою до факту, він є умовою його розуміння. Факт без наративної рамки є просто сирою подією, а наратив перетворює подію на смисл. Цей вимір відповідає некласичній раціональності – увазі до позиції, передумов і дискурсивних умов виробництва знання.

Наративний вимір є тим рівнем, де відбувається головна боротьба в сучасних інформаційних конфліктах. Конфлікт гранд-наративів між Заходом і Росією, боротьба за інтерпретацію війни на міжнародній арені, суперечка за те, що є «агресією» і що є «спеціальною операцією», все це є битвою за наративний вимір (Христокін 2026). П. Помєранцев точно зафіксував: сучасна російська пропаганда не прагне переконати в правдивості свого наративу, вона прагне зруйнувати саму ідею того, що якийсь наратив може бути більш правдивим за інший (Pomerantsev 2019). Це атака на наративний вимір у найрадикальнішій формі.

Ідентифікаційний вимір відповідає за формування колективних уявлень про себе та Іншого. Це рівень, на якому суспільство питає: хто ми? Кому ми довіряємо? Яке місце ми займаємо у світі? Це рівень когнітивних упереджень, стереотипів, образів ворога і союзника, які Д. Канеман і А. Тверськи описали як систематичні відхилення від раціональної обробки інформації (Kahneman, Tversky 1979). Цей вимір відповідає постнекласичній раціональності, увазі до системних умов, за яких формується саме знання.

Ідентифікаційний вимір є тим рівнем, де маніпуляція діє найбільш непомітно. Ніхто не сприймає власні когнітивні упередження як упередження, їх носій перебуває у своєрідній «наративній сліпоті», упередження вони відчуваються як здоровий глузд, як очевидність, як «природний» погляд на речі. Саме тому атака на ідентифікаційний вимір є найглибшою: вона впливає не на те, що люди думають, а на те, як вони думають і кому вони готові довіряти.

Тут важливо зробити принципове уточнення. Чотири виміри епістемічної архітектури є не лише системою захисту від зовнішніх атак. Це умова існування будь-якої спільноти як суб'єкта. Спільнота, яка встановлює факти, будує канали їх поширення, формує наративи і виробляє ідентичність, і тим самим конституює саму себе. Саме тому руйнування будь-якого з цих вимірів є не просто загрозою безпеці, це загроза існуванню «ми» як такого. П. Рікер показав, що наративна ідентичність є не фіксованою властивістю, а постійним процесом самооповіді: «суб'єкт існує тому, що він себе розповідає» (Ricoeur 1990, 147). Перенесене на рівень суспільства, це означає, що суспільство, яке втратило здатність виробляти власні наративи або контролювати канали їх поширення, втрачає не лише «образ себе», воно втрачає здатність до колективної дії.

Поняття «архітектура» використовується для позначення неадитивної системної взаємозалежності чотирьох вимірів. Йдеться не про їх сукупність, а про таку організацію, за якої кожен вимір виконує функцію умови можливості для інших. У цьому сенсі взаємозалежність має характер функціональної комплементарності: порушення одного виміру призводить до каскадної дестабілізації інших, оскільки руйнуються умови їхнього функціонування.

Ця логіка повністю відтворюється на рівні епістемічної архітектури. Верифікаційний вимір без наративного – це набір ізольованих фактів, позбавлених смислу. Будь-яка пропаганда починає саме з того, що визнає певні факти, але вписує їх у власний наративний контекст. Наративний вимір без верифікаційного є чистою риторикою, відірваною від реальності. Комунікаційний вимір без верифікаційного є каналом, що рівноправно транслює правду і маніпуляцію, і в умовах алгоритмічної оптимізації на залучення маніпуляція структурно виграє. Верифікаційний вимір без комунікаційного є правдою, яка не чутна: факти встановлені, але не досягають аудиторії через контрольовані або спотворені канали. Ідентифікаційний вимір без трьох інших – це замкнуте коло групових упереджень, що блокує і верифікацію, і інтерпретацію. А всі чотири разом, без постійного взаємного коригування деградують кожен у власний спосіб.

Чому ми говоримо про архітектуру, а не просто про чотири аспекти безпеки? Тому що між чотирма вимірами існує тип залежності, який не можна описати як просту суму або як ієрархію. Чотири кільця переплетені так, що якщо розрізати будь-яке одне, то всі чотири розпадаються. Це не означає, що вони зливаються, кожне зберігає власну форму і

функцію. Але їх взаємна підтримка є необхідною умовою стабільності кожного.

Що це означає для діагностики інформаційних атак? Стратегія «firehose of falsehood» (шланг неправди), яку К. Пол і М. Меттьюз описали як характерну рису російської пропаганди, є не просто атакою на факти (Paul and Matthews 2016). Це одночасна атака на всі чотири виміри: заперечення верифікованих фактів руйнує верифікаційний вимір; затоплення інформаційного простору суперечливими повідомленнями руйнує комунікаційний вимір; встановлення хибної симетрії наративів руйнує наративний вимір; підрив довіри до інституцій і формування образу «корумпованого Заходу» руйнує ідентифікаційний вимір. Архітектура атакується системно і одночасно.

Це пояснює, чому точкові відповіді – спростування конкретних фейків, захист конкретних наративів, медіаграмотність як індивідуальна навичка є необхідними, але недостатніми. Вони лікують симптоми, не зачіпаючи архітектуру вразливості. Захист епістемічної архітектури вимагає одночасної підтримки всіх чотирьох вимірів.

До недавнього часу різні типи загроз для епістемічної архітектури були більш-менш розмежовані. Одні актори спеціалізувались на фабрикації фактів, інші на наративних маніпуляціях, треті – на впливі на ідентифікаційні процеси. Штучний інтелект змінює цю ситуацію якісно. Він є першим актором, який діє на всіх чотирьох рівнях одночасно, автоматично і в масштабі, недосяжному для людських операторів.

На верифікаційному рівні генеративні моделі здатні виробляти реалістичні тексти, зображення і відео, що є фактичними фальсифікаціями. Синтетичні медіа руйнують базову процедуру верифікації – відрізнення автентичного документу від фабрикації, не через конкретний обман, а через масштабне розмиття самих критеріїв автентичності. Коли будь-яке відео може бути deepfake, суспільство втрачає не просто конкретний документ, воно втрачає довіру до самого класу доказів.

На комунікаційному рівні ШІ-системи здійснюють автоматичне управління інформаційними потоками в масштабі, недоступному людині. Боти і автоматизовані акаунти здатні штучно підвищувати видимість певних повідомлень, створюючи ілюзію масової підтримки і органічного поширення. Рекомендаційні алгоритми, оптимізовані на залучення, систематично підсилюють емоційно заряджений і поляризуючий контент. К. Вордл і Х. Дерахшан зафіксували, що «безлад інформації» є не лише проблемою хибного змісту, а й проблемою архітектури середовища, що визначає, які повідомлення стають вірусними (Wardle, Derakhshan 2017). ШІ перетворює цю архітектурну проблему на інструмент цілеспрямованого впливу.

На наративному рівні генеративні моделі здатні виробляти узгоджені наративи у промисловому масштабі, адаптовані до конкретних аудиторій, мов і культурних кодів. Те, що раніше вимагало армії пропагандистів, тепер виконує модель, оптимізована на «переконливість». Це не просто прискорення -- це зміна природи наративної боротьби. Людина більше не може відрізнити, з ким вона полемізує: з людиною або з алгоритмом.

На рівні ідентичності алгоритми рекомендацій здійснюють систематичну персоналізацію інформаційного середовища кожного користувача. М. Рібейро показав, що алгоритм YouTube систематично веде користувачів у напрямку дедалі більш радикального контенту – це є спроектованим результатом оптимізації на залучення (Ribeiro 2020). Це означає, що ідентифікаційний вимір кожного користувача формується алгоритмом, який не знає і не враховує питань епістемічної безпеки.

Важливо підкреслити: ШІ є актором, але не суб'єктом у філософському сенсі. Він не приймає рішення, не має намірів і не несе відповідальності, але саме це робить його особливо небезпечним. Відповідальність розчиняється в «логіці системи», і саме цей ефект ми позначаємо як деінструменталізацію на рівні постнекласичного виміру. «Алгоритм так вирішив» є формулою, що імітує пояснення і водночас унеможливлює відповідальність.

У зв’язку з цим ми пропонуємо термін, що допоможе посилити спротив проти атак на епістемну безпеку. *Епістемічна резилієнтність* (epistemic resilience) описує здатність суспільства відновлювати цю архітектуру після атаки. Це поняття є важливим, бо ідея абсолютного захисту від атак є ілюзорною. Питання не в тому, чи буде атаковано верифікаційний, комунікаційний, наративний або ідентифікаційний вимір – вони будуть атаковані постійно. Питання в тому, чи здатне суспільство відновити архітектурну цілісність після кожної атаки.

Епістемічна резилієнтність є властивістю одночасно інституцій, суспільства і індивіда, але не редукується до жодного з них окремо. Незалежні організації верифікації фактів утримують верифікаційний вимір. Різноманіття медіаплатформ і прозорість алгоритмів підтримує комунікаційний вимір. Різноманіття наративних платформ підтримує наративний вимір. Критична медіаграмотність як масова практика формує стійкість ідентифікаційного виміру. Але жоден із цих елементів сам по собі не є достатнім: резилієнтність (стійкість) є системною властивістю їх взаємодії.

Гелен Лонгіно описала чотири умови наукової об'єктивності: відкритість до заперечень, публічність стандартів, рівноправна участь у критиці і здатність змінюватися під тиском аргументів (Longino 1990, 76). Ми пропонуємо читати ці умови як опис епістемічної резилієнтності на інституційному рівні. Суспільство є резилієнтним (стійким) тоді, коли ці умови дотримуються стосовно всіх чотирьох вимірів одночасно.

Стратегія «prebunking» (епістемічне щеплення) – попереднє знайомство суспільства з механізмами маніпуляції до їх застосування є одним із найбільш емпірично підтверджених інструментів підвищення епістемічної резилієнтності. М. Бідлстоун, Д. Рузенбек і С. ван дер Ліндер показали, що попереднє «щеплення» проти конкретних маніпулятивних прийомів суттєво знижує їх ефективність (Biddlestone, Roozenbeek, van der

Linden 2022). Це є практичним підтвердженням того, що резилієнтність є не природженою властивістю, а виробленою навичкою.

Чотири питання-критерії операційної оцінки епістемічної резилієнтності суспільства: чи зберігають суспільну вагу незалежні організації верифікації фактів? Чи є прозорими алгоритмічні системи, що формують комунікаційне середовище, і чи існують механізми обмеження маніпулятивних архітектур поширення? Чи є різноманітним наративний простір і чи забезпечена рівноправна участь у ньому різних голосів? Чи аналізуються системні умови, що уможливлюють маніпуляції, і чи є ця аналітика публічно доступною? Позитивна відповідь на всі ці питання характеризує суспільство з розвиненою епістемічною резилієнтністю.

Обговорення

Запропонований концепт епістемічної архітектури безпеки потребує позиціонування відносно кількох споріднених підходів. Найближчим аналогом є підхід Е. Сеґер до «епістемічної безпеки», визначеної як захист систем знання від загроз, що підривають їх цілісність і надійність (Seger et al. 2020). Ця позиція близька до нашої, але вона залишається на рівні опису загроз, не пропонуючи архітектурного концепту, який би пояснював, чому певні атаки є системно ефективними. Наш підхід іде далі: він пояснює механізм вразливості через структуру взаємозалежності чотирьох вимірів.

Дослідники когнітивної війни в рамках НАТО стверджують, що метою сучасних інформаційних кампаній є зміна не лише того, що люди думають, але і того, як вони думають (Du Cluzel 2020). Це точний діагноз, але без філософського інструменту для його концептуалізації. Концепт епістемічної архітектури безпеки дає цей інструмент: «як люди думають» визначається станом усіх чотирьох вимірів архітектури одночасно.

Від С. Левандовського, Ю. Екера і Д. Кука ми знаємо, що в епоху постправди проблема характеризується не браком правдивої інформації, а руйнуванням самих когнітивних механізмів її розпізнавання (Lewandowsky, Ecker, Cook 2017, 353). Це підтверджує нашу тезу про те, що атака на епістемічну архітектуру є ефективнішою, ніж атака на конкретні переконання. Але когнітивна психологія залишається на рівні індивідуального пізнання і не дає інструменту для аналізу системного виміру проблеми.

Від М. Фуко ми запозичуємо розуміння того, що «режими істини» є структурами влади (Foucault 1980). Але М. Фуко описує ці режими як механізми виключення і нормалізації і не дає нормативного критерію того, якою має бути «здорова» епістемічна архітектура. Наш підхід є нормативним: ми стверджуємо, що епістемічна резилієнтність є не лише бажаною, але й необхідною умовою демократичного самовизначення.

Подальшого дослідження потребує питання про позитивний зміст комунікаційного виміру епістемічної архітектури: осмислення не лише того, як суспільство захищається від маніпулятивних каналів, а й того, як воно будує власне автономне комунікаційне середовище, здатне підтримувати верифікаційний і наративний виміри одночасно.

Висновки

Встановлено, що епістемічна архітектура безпеки має чотиривимірну будову (верифікаційний, комунікаційний, наративний та ідентифікаційний виміри), які перебувають у відношенні неадитивної функціональної взаємозалежності. Порушення будь-якого з вимірів має каскадний ефект і дестабілізує систему в цілому, що пояснює обмеженість суто технічних рішень і редукцію атак лише до «фактів» або лише до «наративів».

Також описано специфіку загроз для кожного рівня. Верифікаційний вимір атакується через заперечення фактів, фабрикацію доказів і дискредитацію інституцій верифікації. Комунікаційний вимір атакується через маніпуляцію інформаційними потоками: штучне підвищення видимості дезінформаційного контенту, алгоритмічне пригнічення незручних повідомлень і побудову замкнутих інформаційних бульбашок, які унеможливлюють контакт із альтернативними джерелами. Наративний вимір атакується через встановлення хибної симетрії між несиметричними позиціями і підрив самої ідеї легітимного наративу. Ідентифікаційний вимір атакується через персоналізоване підсилення когнітивних упереджень і руйнування довіри до інституцій. Штучний інтелект є першим актором, що атакує всі чотири виміри одночасно і автоматично, що є якісно новим типом загрози.

Зрештою обґрунтовано поняття епістемічної резилієнтності як здатності суспільства відновлювати архітектурну цілісність після атаки. Ця здатність є системною властивістю взаємодії інституцій, суспільних практик і індивідуальних навичок, і не може бути зведена до жодного з них окремо. Чотири питання-критерії операційної оцінки резилієнтності дозволяють здійснити конкретну оцінку стану безпеки суспільства.

Цей підхід розвиває і поглиблює концепцію структурної кореляції раціональностей, запропоновану в попередній статті, переносячи її з рівня наукового пізнання на рівень суспільного пізнання в умовах кіберсуспільства і показуючи, як ця структура функціонує в умовах системних атак і яким є критерій її збереження.

Список літератури

1. Христокін, Г. В., Верховцева, І. Г., Слюсар, В. Класична, некласична і постнекласична раціональність: структурна кореляція і епістемічна вразливість в умовах кіберсуспільства. Вісник КАІ. Філософія. Культурологія. 2025. № 2 (42). С. 48–56. DOI: https://doi.org/10.18372/2412-2157.42.21000

2. Христокін Г. Наративні та реальні причини російської агресії проти України. Contemporary International Relations: Topical Highlights of Theory and Practice : monograph / ed. by Z. Sharlovych, Y. Voloshyn, N. Vasylyshyna. Łomża : MANS w Łomży, 2024. P. 50–65. DOI: https://doi.org/10.58246/JBBJ2120

3. Auditing radicalization pathways on YouTube / M. H. Ribeiro, R. Ottoni, R. West, V. A. F. Almeida, W. Meira Jr. FAT 2020. P. 131–141. DOI: https://doi.org/10.1145/3351095.3372879.

4. Benkler Y., Faris R., Roberts H. Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics. Oxford: Oxford University Press, 2018. 472 p. https://doi.org/10.1093/oso/9780190923624.001.0001

5. Berger P. L., Luckmann T. The Social Construction of Reality: A Treatise in the Sociology of Knowledge. New York : Doubleday, 1966. 219 p.
6. Biddlestone M., Roozenbeek J., van der Linden S. Psychological inoculation reduces susceptibility to misinformation on social media. Royal Society Open Science. 2022. Vol. 9. Art. 220201. DOI: https://doi.org/10.1098/rsos.220201.
7. Castells M. Communication Power. Oxford : Oxford University Press, 2009. 571 p.
8. Du Cluzel F. Cognitive Warfare. Norfolk, VA : NATO Allied Command Transformation Innovation Hub, 2020. 45 p. https://innovationhub-act.org/wp-content/uploads/2023/12/20210113_CW-Final-v2-.pdf
9. Floridi L. The Fourth Revolution: How the Infosphere Is Reshaping Human Reality. Oxford : Oxford University Press, 2014. 248 p.
10. Foucault M. Power/Knowledge: Selected Interviews and Other Writings, 1972–1977. N. Y. : Pantheon Books, 1980. 270 p.
11. Gillespie T. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. New Haven : Yale University Press, 2018. 296 p. DOI: https://doi.org/10.12987/9780300235029.
12. Habermas J. The Theory of Communicative Action. Vol. 1. Boston: Beacon Press, 1984. 465 p.
13. Kahneman D., Tversky A. Prospect theory: an analysis of decision under risk. Econometrica. 1979. Vol. 47, No. 2. P. 263–291. DOI: https://doi.org/10.2307/1914185
14. Latour B. Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford ; New York : Oxford University Press, 2005. 301 p. DOI: https://doi.org/10.1093/oso/9780199256044.001.0001
15. Lewandowsky S., Ecker U. K. H., Cook J. Beyond misinformation: understanding and coping with the "post-truth" era. Journal of Applied Research in Memory and Cognition. 2017. Vol. 6, No. 4. P. 353–369. DOI: https://doi.org/10.1016/j.jarmac.2017.07.008
16. Longino H. E. Science as Social Knowledge. Princeton: Princeton University Press, 1990. 262 p.
17. Paul C., Matthews M. The Russian "Firehose of Falsehood" Propaganda Model. Santa Monica, CA: RAND Corporation, 2016. DOI: https://doi.org/10.7249/PE198.
18. Pomerantsev P. This Is Not Propaganda: Adventures in the War Against Reality. New York: PublicAffairs, 2019. 256 p.
19. Rid T. Cyber War Will Not Take Place. Oxford: Oxford University Press, 2013. 218 p.
20. Ricoeur P. Oneself as Another. Chicago: University of Chicago Press, 1990. 368 p.
21. Russell S. Human Compatible: Artificial Intelligence and the Problem of Control. New York: Viking, 2019. 352 p.
22. Seger E., Avin S., Pearson G., Briers M., Ó Heigeartaigh S., Bacon H. Tackling threats to informed decision-making in democratic societies: promoting epistemic security in a technologically-advanced world. London : The Alan Turing Institute, 2020. 112 p.
23. Stepin V. S. Theoretical Knowledge: Structure, Historical Evolution. Dordrecht : Springer, 2005. 344 p. (Synthese Library ; vol. 326). DOI: https://doi.org/10.1007/1-4020-3046-0
24. Vosoughi S., Roy D., Aral S. The spread of true and false news online. Science. 2018. Vol. 359, No. 6380. P. 1146–1151. DOI: https://doi.org/10.1126/science.aap9559
25. Wardle C., Derakhshan H. Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking : Council of Europe Report DGI(2017)09. Strasbourg : Council of Europe, 2017. 107 p. URL: https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c.

References

1. Khrystokin, Hennadii, Iryna Verkhovtseva, and Vadym Slyusar. 2025. “Klasychna, neklasychna i postneklasychna ratsionalnist u vymiri tsyfrovoho suverenitetu ta kohnityvnoi bezpeky” [Classical, Non-Classical and Post-Non-Classical Rationality in the Dimension of Digital Sovereignty and Cognitive Security]. *Visnyk Kyivskoho aviatsiinoho universytetu. Seriia: Filosofiia. Kulturolohiia* 2 (42): 41–50.
2. Khrystokin, Hennadii. 2024. “Narratyvni ta realni prychyny rosiiskoi ahresii proty Ukrainy” [Narrative and Real Causes of Russian Aggression against Ukraine]. In *Contemporary International Relations: Topical Highlights of Theory and Practice*, edited by Z. Sharlovych, Y. Voloshyn, and N. Vasylyshyna, 50–65. Łomża: MANS w Łomży. https://doi.org/10.58246/JBBJ2120.
3. Ribeiro, Manoel Horta, Raphael Ottoni, Robert West, Virgílio Almeida, and Wagner Meira Jr. 2020. “Auditing Radicalization Pathways on YouTube.” In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* ’20)*, 131–141. https://doi.org/10.1145/3351095.3372879.
4. Benkler, Yochai, Robert Faris, and Hal Roberts. 2018. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford: Oxford University Press.
5. Berger, Peter, and Thomas Luckmann. 1966. The Social Construction of Reality. New York: Anchor Books.
6. Biddlestone, Mikey, Jon Roozenbeek, and Sander van der Linden. 2022. “Psychological Inoculation Reduces Susceptibility to Misinformation on Social Media.” *Royal Society Open Science* 9: 220201. https://doi.org/10.1098/rsos.220201.
7. Castells, Manuel. 2009. *Communication Power*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199567041.001.0001.
8. Du Cluzel, François. 2020. *Cognitive Warfare*. Norfolk, VA: NATO Allied Command Transformation Innovation Hub.
9. Floridi, Luciano. 2014. *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199606726.001.0001.
10. Foucault, Michel. 1980. *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977*. NY: Pantheon Books.
11. Gillespie, Tarleton. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven: Yale University Press. https://doi.org/10.12987/9780300235029.
12. Habermas, Jürgen. 1984. *The Theory of Communicative Action*. Vol. 1. Boston: Beacon Press.
13. Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47 (2): 263–291. https://doi.org/10.2307/1914185.
14. Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford Univ.Press.
15. Lewandowsky, Stephan, Ullrich Ecker, and John Cook. 2017. “Beyond Misinformation: Understanding and Coping with the Post-Truth Era.” *Journal of Applied Research in Memory and Cognition* 6 (4): 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008.
16. Longino, Helen E. 1990. *Science as Social Knowledge*. Princeton: Princeton University Press.
17. Paul, Christopher, and Miriam Matthews. 2016. *The Russian "Firehose of Falsehood" Propaganda Model*. Santa Monica, CA: RAND Corporation. https://doi.org/10.7249/PE198.
18. Pomerantsev, Peter. 2019. *This Is Not Propaganda: Adventures in the War Against Reality*. New York: PublicAffairs.
19. Rid, Thomas. 2013. *Cyber War Will Not Take Place*. Oxford: Oxford University Press.
20. Ricoeur, Paul. 1990. *Oneself as Another*. Chicago: University of Chicago Press.
21. Russell, Stuart. 2019. *Human Compatible: Artificial Intelligence and the Problem of Control*. New York: Viking.
22. Seger, Elizabeth, Shahar Avin, Gavin Pearson, Mark Briers, Seán Ó hÉigeartaigh, and Helena Bacon. 2020. *Tackling Threats to Informed Decision-Making in Democratic Societies: Promoting Epistemic Security in a Technologically-Advanced World*. London: The Alan Turing Institute. https://www.turing.ac.uk/sites/default/files/2020-10/epistemic-security-report_final.pdf.
23. Stepin, Viacheslav. 2005. *Theoretical Knowledge: Structure, Historical Evolution*. Dordrecht: Springer. https://doi.org/10.1007/1-4020-3673-2.
24. Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018. “The Spread of True and False News Online.” *Science* 359 (6380): 1146–1151. https://doi.org/10.1126/science.aap9559.
25. Wardle, Claire, and Hossein Derakhshan. 2017. *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Strasbourg: Council of Europe.

Hennadii Khrystokin, Iryna Verkhovtseva, Marutian Rena

EPISTEMIC ARCHITECTURE OF SECURITY: PHILOSOPHICAL AND METHODOLOGICAL FOUNDATIONS FOR RESEARCHING INFORMATION AND COGNITIVE SECURITY IN THE CONDITIONS OF CYBER-SOCIETY

Introduction. The question of what exactly is being protected in information security has ceased to be purely rhetorical. Russia's full-scale aggression against Ukraine demonstrated that the most vulnerable object of attack is not infrastructure, but society's ability to distinguish real from constructed. This article develops the concept of structural correlation of rationalities proposed in a preceding study (Khrystokin, Verkhovtseva, and Slyusar 2025). **The aim and tasks.** The article aims to substantiate the concept of epistemic architecture of security as a philosophical and methodological framework for analysing information and cognitive security in cyber society and to show how artificial intelligence transforms threats to this architecture. **Research methods.** The study draws on Stepin's typology of scientific rationality, Longino's social epistemology, Habermas's theory of communicative action, as well as Foucault's concept of regimes of truth, Floridi's concept of the infosphere, and Gillespie's analysis of platform architectures as epistemic actors. **Research results.** The epistemic architecture of security is a four-pole structure whose dimensions – verificational, communicative, narrative, and identificational – stand in a necessary mutual dependency. The verificational dimension addresses the establishment of facts. The communicative dimension concerns the architecture of information flows and their distribution across audiences. The narrative dimension governs the interpretation of events. The identificational dimension encompasses cognitive biases and identity processes that enable manipulation. Artificial intelligence is analysed as an actor that attacks all four dimensions simultaneously. The concept of epistemic resilience is introduced as the society's capacity to restore the integrity of this architecture after an attack. **Conclusions.** Epistemic resilience is a systemic property of the interaction between institutions, social practices, and individual skills. The concept of epistemic architecture of security develops the structural correlation of rationalities proposed in the preceding study, transferring it from the level of scientific cognition to the level of public cognition and providing both a diagnosis of vulnerability and a criterion for its restoration.

Keywords: *epistemic architecture of security, cognitive security, information security, structural correlation of rationalities, epistemic resilience, communicative dimension, narrative security, artificial intelligence, cyber society, de-instrumentalization of rationality.*

Дата першого надходження: 02.02.2026.
Дата прийняття до друку: 31.03.2026.
Дата публікації: 28.05.2026

DOI: https://doi.org/10.18372/2412-2157.1.21228
УДК 2-1:2-9:130.3(045)

СТАНОВЛЕННЯ ТА РОЗВИТОК ЕКЗИСТЕНЦІАЛЬНОЇ ТЕОЛОГІЇ В ІСТОРІОГРАФІЧНОМУ РАКУРСІ

Сергій Шевченко

Національний медичний університет імені О. О. Богомольця
ex.theology@gmail.com | https://orcid.org/0000-0002-9713-3402

Анотація. *Дана розвідка зосереджується на історіографічному та історико-філософському аналізі специфіки співвідношення та взаємовпливу християнської теології та екзистенціалізму. Доводиться, що вивчення методологічної та ідейної кореляції екзистенціалізму і християнської теології на теоретичному релігієзнавчому рівні, і, особливо, яке стосується періоду функціонування та розвитку так званого постекзистенціалістського мислення, лише розпочинається. Зарубіжні дослідження підтверджують доволі явну довільність у визначенні та класифікації головних постулатів релігійних філософів та теологів, як екзистенціалістів, так і екзистенціальних мислителів, що засвідчує наявність методологічних суперечностей в оцінці феномену класичного екзистенціалізму загалом та екзистенціальної теології зокрема (Г. А. Слейта, М. Вестфаля).*

Ключові слова: християнська теологія, екзистенціалізм, екзистенціальна теологія, екзистенціальна діалектика, постекзистенціалістське мислення, філософія релігії

Вступ

Проблема синтезу класичного екзистенціалізму з християнською теологією, зокрема імплементація екзистенціалістських ідей в теологічне мислення американськими теологами Говардом Александером Слейтом та Мерольдом Вестфалем на світоглядному та методологічному рівні, залишається актуальною та малодослідженою в середовищі української історико-філософської та релігієзнавчої думки. Тому вивчення методологічної та ідейної кореляції екзистенціалізму і християнської теології на теоретичному релігієзнавчому рівні, і, особливо, яке стосується періоду функціонування та розвитку так званого постекзистенціалістського мислення, лише розпочинається. Однією з істотних тем такого вивчення є дослідження специфіки становлення та розвитку екзистенціальної теології.

Мета та завдання дослідження

Метою даної розвідки є аналіз особливостей співвідношення та взаємовпливу християнської теології та екзистенціалізму. Завдання полягають у докладному висвітленні широкого масиву, передусім, зарубіжних та українських джерел, присвячених дослідженню творчості Г. А. Слейта та М. Вестфаля, видатних представників екзистенціальної теології.

Методологія дослідження

Методологія дослідження опирається на системний, історичний, діалектичний та аналітичний підходи, що створили умови для об’єктивної інтерпретації складних процесів інтеграції екзистенціалізму з теологією. Компаративний, герменевтичний та контекстуалізаційний методи дозволили проаналізувати, порівняти, співставити та