МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ**

**НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КІБЕРБЕЗПЕКИ ТА ЗАХИСТУ ІНФОРМАЦІЇ
КАФЕДРА УПРАВЛІННЯ КІБЕРБЕЗПЕКОЮ ТА ЗАХИСТОМ ІНФОРМАЦІЇ**

КВАЛІФІКАЦІЙНА РОБОТА

на тему: “МЕТОДИ ЗАХИСТУ ВІД КОГНІТИВНИХ ВИКРИВЛЕНЬ У КІБЕРПРОСТОРІ”

на здобуття освітнього ступеня бакалавра
зі спеціальності 125 Кібербезпека
освітньої програми Управління інформаційною та кібернетичною безпекою

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Ярослав КУЧЕРЯВЕНКО
(підпис) *Ім’я, ПРІЗВИЩЕ здобувача*

Виконав: здобувач вищої освіти гр. УБД-42

Ярослав КУЧЕРЯВЕНКО
Ім’я, ПРІЗВИЩЕ

Керівник: Тетяна КАПЕЛЮШНА
д.е.н., доцент Ім’я, ПРІЗВИЩЕ

Рецензент: Галина Гайдур
д.т.н., професор Ім’я, ПРІЗВИЩЕ

Київ 2026

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИ ТЕХНОЛОГІЙ

Навчально-науковий інститут кібербезпеки та захисту інформації

Кафедра управління кібербезпекою та захистом інформації
Ступінь вищої освіти бакалавр
Спеціальність 125 Кібербезпека
Освітня програма Управління інформаційною та кібернетичною безпекою

ЗАТВЕРДЖУЮ
Завідувач кафедри УКБЗІ
_____ Світлана ЛЕГОМІНОВА
“____”_____2026 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

Кучерявенку Ярославу Володимировичу
(прізвище, ім’я, по батькові здобувача)

1. Тема кваліфікаційної роботи “Методи захисту від когнітивних викривлень у кіберпросторі”,
керівник кваліфікаційної роботи Капелюшна Тетяна, доцент,
(ІПРІЗВИЩЕ, Ім’я,, науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологій від “20“ лютого 2026 р. №51

2. Строк подання кваліфікаційної роботи “12” травня 2026р.
3. Вихідні дані до кваліфікаційної роботи: *когнітивна безпека, інформаційно-психологічні операції, методи та засоби протидії дезінформації, міжнародні стандарти та найкращі практики когнітивної безпеки, наукова та технічна література з когнітивної психології, інформаційної безпеки та гібридної війни.*
4. Перелік питань, які мають бути розроблені:
   4.1. Ознайомитися теоретичні основи природи когнітивних спотворень та концепцію систем мислення в контексті цифрового середовища та когнітивних викривлень, що виникають при споживанні контенту в кіберпросторі.
   4.2. Проаналізувати роль когнітивних викривлень як інструменту реалізації інформаційно-психологічних операцій (ІПСО) та гібридних війн. Оцінити ефективність сучасних методів маніпулятивного впливу та їх вплив на суспільну свідомість.
   4.3. Запропонувати проєкт національної системи когнітивної детекції для виявлення фактологічної інформації та маніпулятивного контенту. Запропонувати метод захисту суспільної свідомості від цілеспрямованих когнітивних атак та сформувати рекомендації щодо його впровадження.

5. Перелік ілюстративного матеріалу: *презентація PowerPoint*

6. Дата видачі завдання “05“ березня 2026 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Етапи кваліфікаційної роботи	Термін виконання етапів роботи	Примітка
1.	Визначення об’єкту, предмету, мети та завдань дослідження.	18.03.2026	
2.	Збір та аналіз літератури.	30.03.2026	
3.	Ознайомитися теоретичні основи природи когнітивних спотворень та концепцію систем мислення в контексті цифрового середовища та когнітивних викривлень, що виникають при споживанні контенту в кіберпросторі.	08.04.2026	
4.	Проаналізувати роль когнітивних викривлень як інструменту реалізації інформаційно-психологічних операцій (ІПСО) та гібридних війн. Оцінити ефективність сучасних методів маніпулятивного впливу та їх вплив на суспільну свідомість.	15.04.2026	
5.	Запропонувати проєкт національної системи когнітивної детекції для виявлення фактологічної інформації та маніпулятивного контенту. Удосконалити метод захисту суспільної свідомості від цілеспрямованих когнітивних атак та сформувати рекомендації щодо його впровадження.	22.04.2026	
6.	Формулювання висновків за результатами проведеного дослідження.	29.04.2026	
7.	Оформлення роботи.	06.05.2026	
8.	Оформлення презентації.	11.05.2026	
9.	Отримання рецензії на роботу.	10.06.2026	
10.	Захист в ЕК.	11.06.2026	

Здобувач вищої освіти _____ Ярослав КУЧЕРЯВЕНКО

(підпис) *(Ім’я, ПРІЗВИЩЕ)*

Керівник

кваліфікаційної роботи _____ Тетяна КАПЕЛЮШНА

(підпис) *(Ім’я, ПРІЗВИЩЕ)*

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КІБЕРБЕЗПЕКИ ТА ЗАХИСТУ ІНФОРМАЦІЇ

ПОДАННЯ
ГОЛОВІ ЕКЗАМЕНАЦІЙНОЇ КОМІСІЇ
ЩОДО ЗАХИСТУ КВАЛІФІКАЦІЙНОЇ РОБОТИ
на здобуття освітнього ступеня бакалавра

Направляється здобувач Кучерявенко Я.В. до захисту кваліфікаційної роботи
(прізвище та ініціали)
за спеціальністю 125 Кібербезпека
(код, найменування спеціальності)
освітньої програми Управління інформаційною та кібернетичною безпекою
(назва)
на тему: "Методи захисту від когнітивних викривлень у кіберпросторі"

Кваліфікаційна робота і рецензія додаються.

Директор ННІЗІ _____ Євгенія ІВАНЧЕНКО
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

Висновок керівника кваліфікаційної роботи

Здобувач Кучерявенко Ярослав Володимирович у кваліфікаційній роботі проаналізував теоретичні основи когнітивних викривлень та особливості їх прояву в сучасному кіберпросторі, дослідив техніки маніпулювання, що експлуатують когнітивні викривлення в інформаційно-психологічних операціях гібридної війни, вивчив існуючі технічні та психологічні методи протидії когнітивним загрозам і розробив практичні рекомендації щодо підвищення когнітивної стійкості.

Кучерявенко Ярослав продемонстрував розуміння актуальної проблеми когнітивної безпеки в умовах гібридної війни, показав високий рівень володіння методами наукового дослідження, здійснив ґрунтовний аналіз наукової літератури, міжнародного досвіду та вітчизняної практики. Робота має чітку структуру, логічну послідовність викладу матеріалу, містить власні узагальнення та практичні пропозиції, які можуть бути використані для підвищення рівня когнітивного захисту населення та персоналу.

Все це дозволяє оцінити кваліфікаційну роботу здобувача Кучерявенка Ярослава позитивно та присвоїти йому кваліфікацію бакалавра з кібербезпеки за освітньою програмою Управління інформаційною та кібернетичною безпекою.

Керівник кваліфікаційної роботи _____ Тетяна КАПЕЛЮШНА
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

"____" _____2026 року

Висновок кафедри про кваліфікаційну роботу

Кваліфікаційна робота розглянута. Здобувач Кучерявенко Я.В. допускається до захисту даної роботи в Екзаменаційній комісії.

Завідувач кафедри
управління кібербезпекою та
захистом інформації _____ Світлана ЛЕГОМІНОВА
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

ВІДГУК РЕЦЕНЗЕНТА

на кваліфікаційну бакалаврську роботу

здобувача вищої освіти КУЧЕРЯВЕНКА Ярослава
на тему “Методи захисту від когнітивних викривлень у кіберпросторі”

Актуальність. У сучасних умовах гібридної війни та стрімкого розвитку цифрових технологій когнітивні викривлення стали одним із найефективніших інструментів інформаційно-психологічних операцій. Алгоритмічна персоналізація, швидке поширення дезінформації та штучний інтелект суттєво посилюють вразливість суспільства до маніпулятивного впливу. Захист від когнітивних загроз вимагає не лише технічних рішень, а й системного розвитку когнітивної стійкості населення.

З огляду на це, дослідження методів захисту від когнітивних викривлень у кіберпросторі є актуальним і важливим науковим завданням у галузі кібербезпеки.

Позитивні сторони.

1. У роботі проведено ґрунтовний аналіз теоретичних основ когнітивних викривлень, їх класифікацію в умовах кіберпростору та особливості прояву в контексті інформаційно-психологічних операцій гібридної війни.

2. Автор здійснив системне дослідження технік маніпулювання, що експлуатують когнітивні викривлення, а також проаналізував ефективність існуючих методів протидії.

3. Кваліфікаційна робота має чітку логічну структуру, містить власні узагальнення, таблиці та класифікації. Матеріал викладено грамотно, з використанням значної кількості джерел (близько 40 найменувань, у тому числі англомовних).

4. За результатами дослідження розроблено практичні рекомендації щодо технічних засобів детекції маніпулятивного контенту та комплексних методів підвищення когнітивної стійкості, які мають прикладне значення.

Недоліки.

Доцільно було б більш детально розглянути практичні аспекти впровадження запропонованих рекомендацій в умовах конкретних організацій або державних установ, а також провести оцінку ефективності окремих інструментів детекції когнітивного маніпулятивного контенту на реальних даних.

Однак зазначені зауваження не впливають на загальну позитивну оцінку кваліфікаційної роботи..

Висновок: Кваліфікаційна робота виконана на належному науково-методичному рівні і заслуговує на високу оцінку, а здобувач КУЧЕРЯВЕНКО Ярослав заслуговує присвоєння кваліфікації бакалавра з кібербезпеки за освітньою програмою Управління інформаційною та кібернетичною безпекою.

Рецензент: Галина ГАЙДУР

підпис Ім’я, ПРІЗВИЩЕ

д.т.н., професор,
завідувач кафедри Систем та
технологій кібербезпеки

РЕФЕРАТ

Кваліфікаційна робота присвячена дослідженню методів захисту від когнітивних викривлень у кіберпросторі. Робота складається зі вступу, трьох розділів, що містять 17 таблиць і 5 рисунків, висновків і списку використаних джерел із 40 найменувань. Загальний обсяг роботи становить 109 аркушів, з яких 6 аркуші займають перелік умовних скорочень і список використаних джерел.

Метою роботи. Розробка та обґрунтування дієвого методу захисту суспільної свідомості від когнітивних викривлень у кіберпросторі для підвищення стійкості користувачів до маніпуляцій та цілеспрямованих інформаційно-психологічних операцій.

Об'єктом дослідження є процес деструктивного інформаційно-психологічного впливу на свідомість користувачів у кіберпросторі.

Предмет дослідження **–** методи та засоби захисту від когнітивних викривлень у кіберпросторі.

Методи дослідження. Для вирішення поставлених завдань у роботі використані методи теоретичного аналізу та синтезу, систематизації, класифікації, порівняльного аналізу, структурно-функціонального підходу, а також вивчення наукової літератури та нормативних документів.

Як результат у роботі проаналізовано теоретичні основи когнітивних викривлень та особливості їх прояву в кіберпросторі, досліджено техніки маніпулювання, що експлуатують когнітивні викривлення в інформаційно-психологічних операціях гібридної війни, вивчено існуючі технічні та психологічні методи протидії когнітивним загрозам і розроблено практичні рекомендації щодо підвищення когнітивної стійкості.

Галузь застосування. Результати роботи можуть бути використані в системах забезпечення національної когнітивної безпеки, при розробці програм підвищення обізнаності населення та персоналу щодо протидії дезінформації, а також у діяльності державних органів, що відповідають за інформаційну та кібербезпеку.

Ключові слова: КОГНІТИВНІ ВИКРИВЛЕННЯ, КОГНІТИВНА БЕЗПЕКА, ІНФОРМАЦІЙНО-ПСИХОЛОГІЧНІ ОПЕРАЦІЇ, ГІБРИДНА ВІЙНА, ЗАХИСТ ВІД ДЕЗІНФОРМАЦІЇ, КОГНІТИВНА СТІЙКІСТЬ, КІБЕРПРОСТІР.

ABSTRACT

The qualification work is devoted to the study of methods for protection against cognitive biases in cyberspace. The work consists of an introduction, three chapters containing 17 tables and 5 images, conclusions, and the list of references with 40 sources. The total volume of the work is 109 pages, of which 6 pages are occupied by the list of abbreviations and the list of references.

The purpose of the study is to develop theoretical foundations and practical recommendations for protection against cognitive biases in cyberspace under the conditions of modern information-psychological operations.

The object the study is the methods and means of protection against cognitive biases in cyberspace.

The subject of the study is the peculiarities of applying technologies of information security awareness and training for personnel.

Research methods. To solve the set tasks, the work used methods of theoretical analysis and synthesis, systematization, classification, comparative analysis, structural-functional approach, as well as the study of scientific literature and normative documents.

As a result, the work analyzed the theoretical foundations of cognitive biases and the peculiarities of their manifestation in cyberspace, investigated manipulation techniques that exploit cognitive biases in information-psychological operations of hybrid warfare, studied existing technical and psychological methods of countering cognitive threats, and developed practical recommendations for increasing cognitive resilience.

Field of application. The results of the work can be used in national cognitive security systems, in the development of programs to increase public and personnel awareness on countering disinformation, as well as in the activities of state bodies responsible for information and cybersecurity.

Keywords**:** COGNITIVE BIASES, COGNITIVE SECURITY, INFORMATION-PSYCHOLOGICAL OPERATIONS, HYBRID WARFARE,

PROTECTION AGAINST DISINFORMATION, COGNITIVE RESILIENCE, CYBERSPACE.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ

КБ	Когнітивна безпека
ІПСО	Інформаційно-психологічні операції
КВ	Когнітивні викривлення
ШІ	Штучний інтелект
Deepfake	Діпфейк
Система 1	Швидка (інтуїтивна) система мислення
Система 2	Повільна (аналітична) система мислення
NCTDS	National Cognitive Threat Detection System (Національна система детекції когнітивних загроз)
Backfire effect	Ефект бумеранга
Inoculation Theory	Теорія психологічної імунізації
Confirmation Bias	Підтверджувальне упередження
Availability Heuristic	Евристика доступності

ВСТУП

Актуальність теми. Сучасний етап розвитку інформаційного суспільства характеризується стрімкою цифровою трансформацією, яка радикально змінила способи виробництва, поширення та споживання інформації. У кіберпросторі, де щосекунди генеруються мільйони одиниць контенту, когнітивні викривлення набувають нового, загрозливого виміру. Вони перетворюються з індивідуальних психологічних особливостей на потужний інструмент масового впливу в умовах інформаційно-психологічних операцій (ІПСО) та гібридної війни.

З огляду на зазначене дослідження методів захисту від когнітивних викривлень у кіберпросторі є актуальним науковим завданням.

Метою роботи. Розробка та обґрунтування дієвого методу захисту суспільної свідомості від когнітивних викривлень у кіберпросторі для підвищення стійкості користувачів до маніпуляцій та цілеспрямованих інформаційно-психологічних операцій.

Об'єктом дослідження є процес деструктивного інформаційно-психологічного впливу на свідомість користувачів у кіберпросторі.

Предмет дослідження **–** методи та засоби захисту від когнітивних викривлень у кіберпросторі.

Для досягнення цієї мети в роботі необхідно виконати наступні **завдання**:

Ознайомитися теоретичні основи природи когнітивних спотворень та концепцію систем мислення в контексті цифрового середовища та когнітивних викривлень, що виникають при споживанні контенту в кіберпросторі.

Проаналізувати роль когнітивних викривлень як інструменту реалізації інформаційно-психологічних операцій (ІПСО) та гібридних війн. Оцінити ефективність сучасних методів маніпулятивного впливу та їх вплив на суспільну свідомість.

Запропонувати проєкт національної системи когнітивної детекції для виявлення фактологічної інформації та маніпулятивного контенту. Запропонувати метод захисту суспільної свідомості від цілеспрямованих

когнітивних атак та сформувати рекомендації щодо його впровадження.

Методи дослідження. Для вирішення означеного вище наукового завдання в роботі використані методи теоретичного аналізу і синтезу, систематизація, класифікація, порівняльний аналіз, структурно-функціональний підхід, вивчення наукової літератури та нормативних документів.

Практичне значення одержаних результатів. Розроблені рекомендації можуть бути використані в системах забезпечення національної когнітивної безпеки, при розробці програм підвищення обізнаності населення та персоналу, у діяльності державних органів, що відповідають за інформаційну та кібербезпеку, а також у освітньому процесі.

Апробація результатів кваліфікаційної роботи відбулася на Всеукраїнській науково-практичній конференції “Стратегії кіберстійкості: управління ризиками та безперервність бізнесу” 25 лютого 2026 року.

Розділ 1 ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ДОСЛІДЖЕННЯ КОГНІТИВНИХ ВИКРИВЛЕНЬ

1.1. Концепція систем мислення та природа когнітивних спотворень

Сучасний етап розвитку суспільства характеризується стрімкою цифровою трансформацією, яка радикально змінила способи виробництва, поширення та споживання інформації. У кіберпросторі, де щосекунди генеруються мільйони одиниць контенту, когнітивні викривлення (cognitive biases) набувають нового, загрозливого вектору взаємодії. Вони стають не лише індивідуальними психологічними особливостями, а й потужним інструментом масового впливу в умовах інформаційних операцій та гібридної війни. Актуальність теми зумовлена тим, що людське мислення, еволюційно пристосоване до обмеженого обсягу інформації, виявляється вразливим перед викликами цифрового середовища — алгоритмічною персоналізацією, емоційним зараженням та швидкістю поширення маніпулятивного контенту.

За даними Reuters Institute Digital News Report 2025, понад 58% респондентів у глобальному масштабі висловлюють стурбованість своєю здатністю відрізняти правду від фейків в онлайн-новинах [1]. У багатьох країнах цей показник сягає 70% і вище. Згідно зі звітом ООН Global Risks Report 2024–2025, дезінформація визнана одним з найсерйозніших глобальних ризиків, до якого країни найменше готові [2]. Медіанна оцінка у 25 країнах за дослідженням Pew Research Center (2025) показує, що 72% дорослих вважають поширення неправдивої інформації онлайн серйозною загрозою для своєї країни [3].

У контексті гібридної війни, яку Російська Федерація веде проти України, когнітивні викривлення стають ключовим елементом інформаційно-психологічних операцій. Дезінформація активно використовується для маніпулювання громадською думкою, підриву довіри до державних інститутів та впливу на прийняття рішень як на індивідуальному, так і на колективному рівнях. Дослідження показують, що когнітивні спотворення, такі як підтверджувальне упередження (confirmation bias) та Евристика доступності

(availability heuristic), суттєво посилюють сприйнятливість населення до маніпулятивного контенту, що призводить до спотворення реальності та ірраціональних рішень у критичних ситуаціях [4].

Необхідність теоретичного осмислення проблеми когнітивних викривлень зумовлена тим, що без глибокого розуміння механізмів їхньої дії неможливо розробити ефективні методи захисту. Класичні роботи Д. Канемана, А. Тверскі та Д. Аріелі демонструють, що когнітивні спотворення є системними особливостями людського мислення, а не випадковими помилками. Саме тому теоретичне підґрунтя стає фундаментом для практичних рекомендацій щодо підвищення когнітивної стійкості в цифровому середовищі. Без такого фундаменту будь-які технічні чи освітні заходи матимуть фрагментарний характер [5].

Окремої уваги заслуговує стан когнітивної безпеки в українській освітній системі. На сьогодні в більшості навчальних програм вищих закладів освіти відсутні системні курси з медіаграмотності, психології сприйняття інформації та протидії когнітивним маніпуляціям. Існуючі дисципліни з інформаційної безпеки здебільшого зосереджені на технічних аспектах (кібербезпека, захист даних), тоді як когнітивно-психологічний вимір залишається поза увагою. Така прогалина робить молоде покоління особливо вразливим до впливу сучасних інформаційних загроз і вимагає негайної корекції освітніх стандартів.

Актуальність дослідження когнітивних викривлень у кіберпросторі визначається не лише теоретичною ознайомленістю, але й нагальними практичними потребами національної безпеки України в умовах гібридної війни.

Розуміння механізмів людського мислення пройшло тривалий шлях еволюції в межах психологічної науки. Від античних філософських уявлень психологія поступово переходила до емпіричного вивчення когнітивних процесів. Однак найбільш радикальні трансформації відбулися у XX столітті.

Домінування біхевіоризму в першій половині XX століттясуттєво обмежило вивчення внутрішніх психічних процесів. Біхевіористи розглядали мислення виключно через призму стимул-реакції, відкидаючи свідомість як

ненаукову категорію. Такий підхід, хоча й дозволив накопичити значний експериментальний матеріал, не міг пояснити складні процеси прийняття рішень, суджень та помилок у реальному житті.

Поворотним моментом став «когнітивний поворот» (cognitive revolution) кінця 1950-х – початку 1960-х років. Завдяки роботам Дж. Міллера, У. Найссера, Г. Саймона та інших дослідників психологія повернулася до вивчення внутрішніх ментальних процесів [6]. Було визнано, що людина не є пасивним реагентом на стимули, а активним обробником інформації, який використовує обмежені когнітивні ресурси. Саме в цей період закладаються основи для розуміння мислення як системи обробки інформації.

Особливо значущий внесок у розвиток концепції систем мислення зробили дослідження 1970–1980-х років, проведені Деніелом Канеманом та Амосом Тверскі . Їхня спільна робота «Судження за умов невизначеності: евристики та упередження» стала фундаментальною [7]. Автори показали, що люди в умовах невизначеності та обмеженого часу покладаються на евристики — спрощені правила судження, які в більшості випадків є корисними, але систематично призводять до помилок і когнітивних спотворень.

Подальший розвиток цих ідей знайшов відображення у збірці «Судження за умов невизначеності: евристики та упередження» де було систематизовано емпіричні дані про механізми людських суджень [8]. Ці попередні дослідження заклали основу для розуміння мислення як розрізненого процесу. Остаточного теоретичного оформлення ця ідея набула у фундаментальній монографії Деніела Канемана «Мислення швидке й повільне». У книзі автор чітко розмежовує дві системи мислення: швидку, інтуїтивну Систему 1 та повільну, аналітичну Систему [9].

Розуміння історичного розвитку концепції систем мислення має принципове значення для аналізу сучасних цифрових викликів. Механізми, що сформувалися в умовах обмеженої інформації та відносно стабільного середовища, опинилися в умовах інформаційного перевантаження кіберпростору. Саме тому теоретичне осмислення еволюції поглядів на людське

мислення від біхевіоризму через когнітивну революцію до сучасних двоїстих моделей є необхідною передумовою для вивчення природи когнітивних спотворень у цифровому середовищі та розробки ефективних методів захисту.

Історичний аналіз даних щодо дезінформації вказує, що когнітивні спотворення є не випадковими помилками, а закономірним наслідком еволюційно сформованих механізмів мислення, які потребують глибокого теоретичного осмислення в контексті сучасних інформаційних загроз.

Двоїста модель мислення Деніела Канемана стала однією з найвпливовіших концепцій сучасної когнітивної психології. Лауреат Нобелівської премії з економіки запропонував розглядати людське мислення як взаємодію двох різних систем — Системи 1 і Системи 2, — які суттєво відрізняються за швидкістю, рівнем зусиль та роллю в обробці інформації й прийнятті рішень [10].

Система 1 — це швидка, автоматична, інтуїтивна система мислення. Вона працює постійно, без зусиль і свідомого контролю. Її діяльність характеризується високою швидкістю, асоціативністю, емоційністю та залежністю від евристик. Система 1 формує перші враження, генерує інтуїтивні судження та реагує на стимули практично миттєво. Її сильною стороною є ефективність в умовах обмеженого часу та ресурсів: вона дозволяє людині швидко орієнтуватися в знайомому середовищі, розпізнавати шаблони та реагувати на загрози. Однак Система 1 схильна до помилок, оскільки покладається на спрощені правила (евристики) і сильно піддається впливу емоцій, контексту та когнітивних спотворень.

Система 2, навпаки, є повільною, енергозатратною, аналітичною та свідомою. Вона залучається, коли завдання вимагає уваги, логічного аналізу, складних розрахунків або подолання інтуїтивних імпульсів. Система 2 відповідає за самоконтроль, стратегічне мислення, перевірку гіпотез і прийняття обґрунтованих рішень. Її головна перевага — точність і раціональність. Водночас ця система є «ледачою»: вона споживає значну кількість когнітивних ресурсів і часто уникає активації, якщо Система 1 вже надала прийнятне рішення.

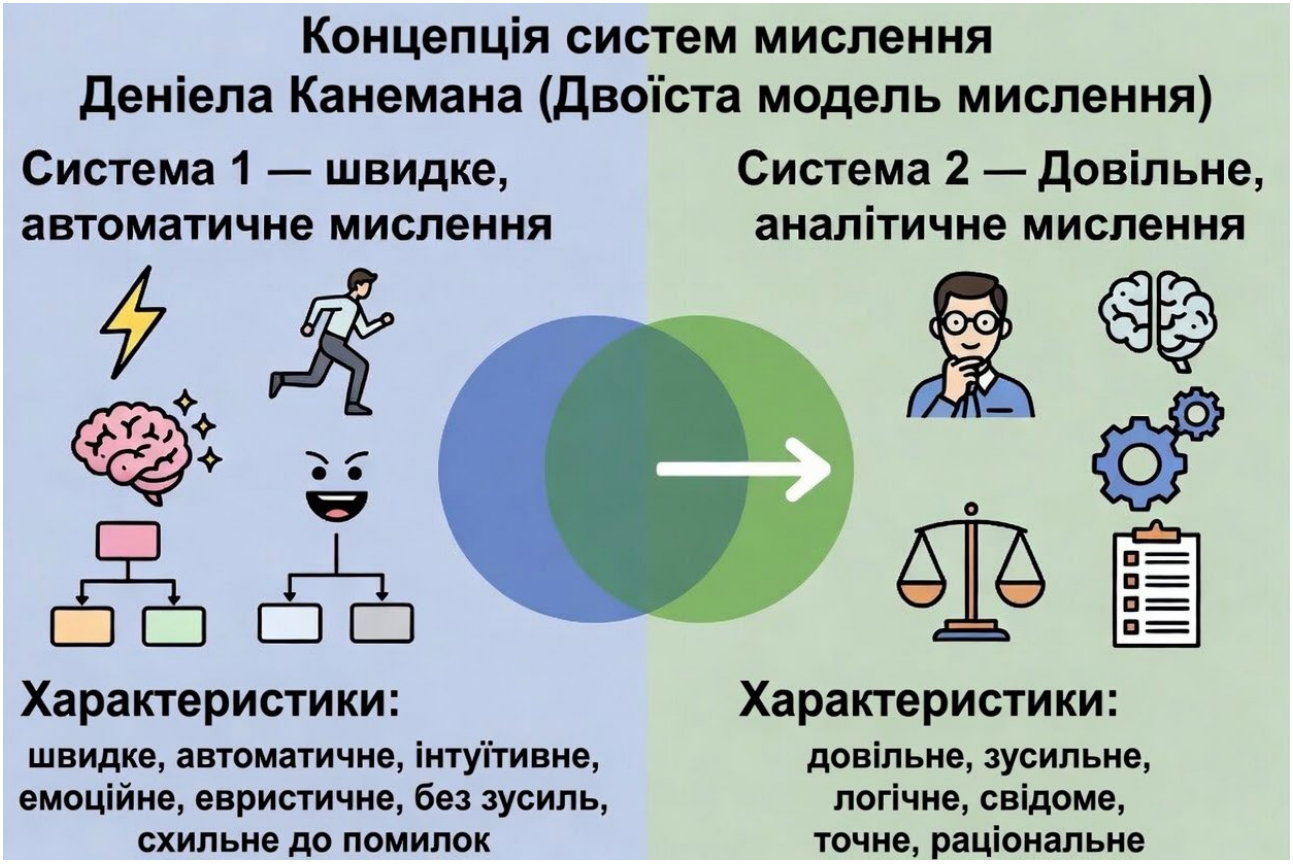

Рис. 1.1. Концепція систем мислення Деніела Канемана

Між двома системами існує постійна взаємодія. Система 1 постійно генерує враження, почуття та інтуїтивні судження, які Система 2 може прийняти, модифікувати або відхилити. У більшості повсякденних ситуацій рішення приймаються саме Системою 1, а Система 2 виконує роль «наглядача», який втручається лише в разі потреби. Однак типовим конфліктом є ситуація, коли інтуїтивне судження Системи 1 суперечить раціональному аналізу. У таких випадках часто перемагає Система 1, особливо за умов когнітивного навантаження, втоми чи емоційного збудження.

Класичні дослідження Канемана та Тверскі яскраво ілюструють роботу цих систем. Одним із найвідоміших прикладів є задача з «Ліндою» Учасникам експерименту повідомляли, що Лінда — 31-річна жінка, яка вивчала філософію, активно цікавилася соціальними питаннями та брала участь у антиядерних демонстраціях. Потім їх просили оцінити ймовірність того, що Лінда є «банківською касиркою» або «банківською касиркою та активісткою

феміністичного руху». Більшість респондентів вважали друге твердження ймовірнішим, хоча це порушує закони логіки (кон'юнкція не може бути ймовірнішою за окрему подію). Це явище, назване «помилкою кон'юнкції», демонструє, як Система 1, спираючись на репрезентативність, ігнорує базові статистичні правила, які могла б застосувати Система 2.

Іншим показовим прикладом є ефект якоря (anchoring). Коли учасників просили оцінити відсоток африканських країн в ООН після того, як вони бачили довільне число (наприклад, 10 або 65), їхні оцінки сильно залежали від початкового «якоря». Це свідчить про те, що Система 1 автоматично використовує першу доступну інформацію, а Система 2 часто не докладає достатніх зусиль для її корекції.

Розуміння взаємодії Системи 1 і Системи 2 дозволяє пояснити, чому когнітивні спотворення є не випадковими помилками, а закономірним наслідком функціонування людського мислення. Саме ця модель слугує теоретичним фундаментом для аналізу природи когнітивних спотворень та їхнього посилення в умовах кіберпростору, що буде розглянуто в наступних підрозділах.

Під когнітивними спотвореннями розуміють стійкі відхилення у процесі сприйняття, інтерпретації та оцінювання інформації, які виникають через особливості функціонування когнітивної системи людини. На відміну від випадкових помилок, когнітивні спотворення є передбачуваними та відтворюваними в певних умовах. Вони безпосередньо пов'язані з використанням евристик — спрощених стратегій мислення, які дозволяють швидко приймати рішення за умов обмеженої інформації та часу.

Евристики — це ментальні «скорочення», що еволюційно склалися для забезпечення ефективної адаптації в умовах ресурсної обмеженості.

Когнітивні спотворення, у свою чергу, є негативними наслідками застосування цих евристик у ситуаціях, де вони не відповідають вимогам раціонального аналізу. Д. Канеман підкреслює, що такі спотворення виникають не через брак інтелекту, а через фундаментальні особливості архітектури мислення

Д. Аріелі вводить поняття «передбачуваної ірраціональності»), стверджуючи, що ірраціональна поведінка людини є не хаотичною, а систематичною та прогнозованою. Він демонструє, що багато рішень, які здаються ірраціональними з точки зору класичної економічної теорії, насправді є закономірним результатом роботи психологічних механізмів

Механізми виникнення когнітивних спотворень мають декілька фундаментальних причин:

1. Економія когнітивних ресурсів. Людський мозок прагне мінімізувати витрати енергії. Система 1, будучи швидкою та автоматичною, постійно використовує евристики замість енергоємного аналізу Системи 2. Така стратегія є ефективною в простих умовах, але призводить до помилок у складних або невизначених ситуаціях

2. Вплив емоцій. Емоційний стан суттєво впливає на процес судження. Позитивні емоції можуть посилювати оптимізм і недооцінку ризиків, тоді як негативні — сприяють песимістичним упередженням. Емоції часто передує когнітивній оцінці та суттєво її спотворюють

3. Контекстуальне сприйняття. Людина сприймає інформацію не ізольовано, а в певному контексті, який сильно впливає на інтерпретацію. Ефект якоря, ефект обрамлення та інші спотворення демонструють, як зміна контекстуальної подачі інформації може радикально змінити рішення при однакових факт

Сучасні дослідження підтверджують, що когнітивні спотворення є універсальними для всіх людей незалежно від рівня освіти чи інтелекту. Вони являють собою вбудовані особливості когнітивної архітектури, сформовані в процесі еволюції для виживання в умовах невизначеності та дефіциту інформації

Природа когнітивних спотворень полягає в їхній системності та еволюційній обумовленості. Вони є невід'ємною частиною функціонування систем мислення і проявляються особливо яскраво в умовах інформаційного перевантаження. Глибоке розуміння цих механізмів створює необхідне

теоретичне підґрунтя для систематизації основних типів когнітивних викривлень.

Одним з найпоширеніших підходів до систематизації когнітивних викривлень є їх класифікація за функціональними групами, залежно від того, на якому етапі процесу мислення вони виникають і яку адаптивну функцію виконували в еволюційному минулому.

Когнітивні спотворення проявляються у вигляді стійких систематичних відхилень у процесі формування суджень і прийняття рішень. Вони виникають через взаємодію Системи 1 і Системи 2 та використання евристичних стратегій. Нижче систематизовано найважливіші типи когнітивних викривлень, розкрито їхні психологічні механізми, наведено класичні експерименти та приклади з подальшим аналізом їхньої дії в умовах звичайного та цифрового середовища

Таблиця 1.1.

Основні когнітивні спотворення: механізми та прояви в реальному й цифровому середовищі

№	Когнітивне спотворення	Психологічний механізм	Прояви в звичайному житті	Прояви в цифровому середовищі
1	Підтверджувальне упередження	Система 1 автоматично віддає перевагу інформації, яка узгоджується з наявними переконаннями, зменшуючи когнітивний дисонанс	Читання лише тих новин і думок, які відповідають власним політичним чи релігійним поглядам	Алгоритмічні рекомендації соцмереж і пошукових систем створюють «інформаційні бульбашки» та посилюють поляризацію суспільства
2	Ефект якоря	Недостатня корекція Системою 2 після встановлення початкового значення (якоря) Системою 1	Переговори, маркетинг (перша оголошена ціна сильно впливає на фінальну)	Перші заголовки новин, вірусні пости або початкова цифра в коментарях суттєво спотворюють сприйняття всієї теми

Продовження табл. 1.1

3	Евристика доступності	Оцінка ймовірності за легкістю згадування прикладів, яку Система 1 сприймає як ознаку частоти	Переоцінка ризиків авіакатастроф після гучних новин	Сенсаційні та емоційно заряджені новини в соцмережах значно спотворюють сприйняття реальних ризиків і тенденцій
4	Евристика репрезентативності	Заміна статистичного аналізу простим порівнянням зі стереотипом або прототипом	Стереотипізація людей за зовнішністю, професією чи національністю	Поширення «типових історій» і вірусних кейсів замість статистики; швидке формування суджень за яскравими прикладами
5	Упередження надмірної впевненості	Ілюзія контролю та недостатня метакогнітивна рефлексія Системи 2	Переоцінка власних прогнозів у фінансах, бізнесі, медицині	Впевнене поширення неперевіреної інформації, фейків і конспірологічних теорій
6	Емоційні упередження	Прямий вплив емоційного ставлення на оцінку ризиків і переваг	Рішення про інвестиції, вибір партнера чи продукту під впливом емоцій	Емоційно заряджений контент (хайп, страх, обурення) домінує в стрічках і сильно впливає на громадську думку

У звичайних умовах ці когнітивні спотворення забезпечують швидку адаптацію та економію ресурсів. Однак в умовах кіберпростору, що характеризується інформаційним перевантаженням, алгоритмічною персоналізацією та високою емоційною напругою, вони стають особливо небезпечними. Надлишок інформації активізує Систему 1, зменшуючи залучення Системи 2, що призводить до масового поширення спотворених уявлень і полегшує маніпулятивний вплив

Сучасний кіберпростір створює принципово нове середовище функціонування систем мислення, яке радикально відрізняється від умов, в яких формувалися когнітивні механізми людини протягом еволюції. Еволюційно сформовані евристики та Система 1, оптимізовані для швидкого реагування в умовах обмеженої інформації та відносно стабільного соціального середовища, опинилися в умовах інформаційного перевантаження

Одним із ключових викликів є алгоритмічна персоналізація контенту в соціальних мережах. Рекомендаційні системи платформ аналізують попередню поведінку користувача і постійно пропонують матеріали, які відповідають його існуючим упередженням. Це суттєво посилює підтверджувальне упередження , оскільки Система 1 отримує постійне підкріплення вже сформованих уявлень, а Система 2 рідко залучається для критичної перевірки. У результаті формується «інформаційний бульбашковий ефект», коли людина живе в штучно створеному інформаційному середовищі, що спотворює сприйняття реальності.

Другим важливим фактором є швидкість поширення інформації. Якщо раніше для поширення новин були потрібні години або дні, то сьогодні вірусний контент може охопити мільйони користувачів за лічені хвилини. Така динаміка активно експлуатує евристику доступності: чим частіше і яскравіше повідомлення з’являється в стрічці, тим вищою здається його ймовірність та значущість, незалежно від реальної правдивості. Швидкість не залишає часу для активації Системи 2, що призводить до імпульсивного поширення та сприйняття інформації.

Значно посилює вразливість когнітивної системи штучний інтелект. Сучасні генеративні моделі здатні створювати високоякісний персоналізований контент, який ідеально адаптується під емоційний стан і когнітивні упередження конкретного користувача. Це робить маніпуляцію контекстом) та емоційним забарвленням надзвичайно ефективною. У цифровому середовищі емоційно заряджений контент (особливо негативний) поширюється значно швидше, активізуючи афективну евристику і знижуючи критичність мислення.

Особливу небезпеку становить емоційне зараження в онлайн-просторі. Емоції, як і віруси, швидко передаються через соціальні мережі, посилюючи колективні когнітивні спотворення. Позитивні чи негативні емоційні тригери, поширювані через яскраві візуальні матеріали та короткі повідомлення, змушують Систему 1 домінувати над аналітичними процесами. У таких умовах людина стає більш схильною до когнітивних упереджень та евристики репрезентативності, сприймаючи яскраві, емоційно насичені приклади як типові.

Сучасний кіберпростір створює ситуацію радикальної невідповідності між еволюційно сформованими механізмами мислення та реаліями цифрового середовища. Класичні когнітивні спотворення, описані Канеманом, Тверскі та Аріелі, не лише зберігаються, але й значно посилюються через архітектуру цифрових платформ. Алгоритми, швидкість і персоналізація роблять Систему 1 постійно перевантаженою, а Систему 2 — недостатньо залученою. Розуміння цих викликів є ключовим для подальшого аналізу специфіки когнітивних викривлень у контексті споживання контенту та розробки

Підсумовуючи, можна стверджувати, що когнітивні спотворення не є випадковими помилками чи наслідком низького інтелекту, а являють собою невід'ємну, системну особливість людського мислення. Швидка, автоматична та інтуїтивна Система 1, яка постійно генерує судження на основі евристик, і повільна, зусильна, аналітична Система 2 утворюють єдину когнітивну архітектуру, в якій перша система домінує в більшості повсякденних рішень.

Когнітивні спотворення є природною складовою людського мислення, яка забезпечувала виживання в умовах обмеженої інформації, але стає джерелом вразливості в умовах цифрової реальності. Глибоке теоретичне осмислення концепції систем мислення та механізмів когнітивних спотворень створює необхідне методологічне підґрунтя для їх подальшого дослідження.

Саме це розуміння дозволяє перейти до аналізу специфіки прояву когнітивних викривлень у контексті споживання контенту в кіберпросторі , де класичні механізми набувають нових форм під впливом алгоритмів, персоналізації та інформаційного перевантаження. У свою чергу, це слугуватиме основою для вивчення соціально-психологічного впливу таких викривлень на суспільство в умовах гібридних загроз

1.2. Класифікація когнітивних викривлень у контексті споживання контенту в кіберпросторі

Сучасний кіберпростір радикально трансформував процеси споживання інформації, зробивши когнітивні викривлення одним із ключових факторів вразливості як окремої людини, так і суспільства в цілому. Якщо загальна теорія когнітивних спотворень добре розроблена в класичній психології, то їхня специфіка в умовах цифрового середовища потребує окремої систематизації та класифікації. Загальні класифікації, орієнтовані на традиційне середовище, не враховують у повній мірі вплив алгоритмічних систем, інформаційного перевантаження, вірусності контенту та персоналізації, що суттєво змінює механізми прояву та силу дії когнітивних викривлень.

Актуальність розробки спеціалізованої класифікації зумовлена стрімким зростанням впливу цифрових платформ на формування громадської думки. За даними досліджень, сучасні користувачі щодня стикаються з тисячами інформаційних повідомлень, більшість з яких обробляється автоматично, з мінімальною участю критичного мислення. У таких умовах когнітивні викривлення стають не просто індивідуальними психологічними особливостями, а системним фактором масового маніпулювання свідомістю.

Особливо гостро ця проблема проявляється в умовах гібридної війни. Як зазначає І.Р. Малик, у сучасній когнітивній війні людська свідомість перетворилася на головне поле бою, де когнітивні викривлення використовуються як високоефективна зброя для дестабілізації суспільства [11]. В умовах інформаційно-психологічних операцій проти України класифікація когнітивних викривлень набуває не лише теоретичного, а й прикладного значення для національної безпеки.

Дослідження Ward et al. (2025) демонструє, що цифрове середовище значно посилює вплив когнітивних упереджень на процес прийняття рішень порівняно з офлайн-середовищем [12]. У свою чергу, Мамчин М.М. (2023) підкреслює, що когнітивні викривлення в умовах надлишку інформації суттєво спотворюють

раціональне сприйняття дійсності, що особливо небезпечно в періоди соціально-політичної нестабільності [13].

Розробка спеціалізованої класифікації когнітивних викривлень у контексті споживання контенту в кіберпросторі є актуальним науковим завданням. Вона дозволить не лише систематизувати знання про механізми вразливості людини в цифровому середовищі, але й створити практичне підґрунтя для підвищення когнітивної стійкості українського суспільства в умовах постійних інформаційних загроз.

Класифікація когнітивних викривлень є важливим методологічним інструментом, який дозволяє систематизувати різноманітні прояви упереджень мислення та зрозуміти механізми їхньої дії в конкретних умовах. У сучасній психологічній науці існує кілька основних підходів до класифікації когнітивних спотворень, які відрізняються критеріями групування.

1. Перший підхід — класифікація за механізмами виникнення. У його межах спотворення поділяються на групи залежно від того, який когнітивний процес вони порушують: сприйняття, інтерпретацію, запам’ятовування чи оцінювання інформації. Найпоширенішою є класифікація за типом евристик (евристика доступності, репрезентативності, якоря тощо), запропонована Тверскі та Канеманом і розвинена в подальших роботах.

2. Другий підхід — класифікація за сферою впливу. Тут спотворення групуються залежно від того, на яку сферу діяльності людини вони найбільше впливають: прийняття рішень, соціальну взаємодію, економічну поведінку чи сприйняття ризиків.

3. Третій підхід — класифікація за ступенем усвідомлення. У цьому випадку розрізняють свідомі (усвідомлювані) та несвідомі (автоматичні) упередження. Більшість когнітивних викривлень належать до другої групи, оскільки вони активуються Системою 1 автоматично, без участі критичного аналізу.

У контексті цифрового середовища ці традиційні класифікації потребують суттєвої адаптації. Сучасні дослідження підкреслюють, що кіберпростір не лише

посилює окремі спотворення, але й створює нові комбіновані форми їх прояву. З огляду на це, цифрове середовище створює унікальні умови, де когнітивні викривлення взаємодіють з алгоритмічними системами, що призводить до виникнення «цифрових гібридних упереджень», які важко пояснити традиційними класифікаціями.

Сучасні практико-орієнтовані дослідження також акцентують увагу на цифровій специфіці. Згідно з матеріалами «Identify biases in the digital world» (TSW, 2023), у онлайн-середовищі особливо небезпечними стають комбінації підтверджувального упередження з ефектом інформаційних бульбашок [14].

Враховуючи вищезазначене, в результаті дослідження обрано багатофакторну (комбіновану) класифікацію когнітивних викривлень. Такий підхід дозволяє поєднати:

- механізм виникнення (когнітивний, емоційний, соціальний);
- умови прояву в цифровому середовищі (алгоритмічна персоналізація, вірусність, тип контенту);
- ступінь впливу на споживання інформації.

Багатофакторна класифікація є найбільш життєздатною для аналізу кіберпростору, оскільки враховує динамічний характер взаємодії людини з цифровим контентом і дозволяє виявити найбільш небезпечні комбінації спотворень в умовах інформаційних операцій. Цей підхід не суперечить класичним теоріям, але розвиває їх відповідно до реалій сучасного цифрового середовища.

Теоретичний аналіз існуючих підходів до класифікації дозволяє перейти до розробки спеціалізованої класифікації когнітивних викривлень, адаптованої до контексту споживання контенту в кіберпросторі

Для системного аналізу особливостей прояву когнітивних викривлень у процесі споживання контенту в кіберпросторі пропонується багатофакторна класифікація за домінуючим механізмом впливу на сприйняття інформації. Такий підхід дозволяє врахувати специфіку цифрового середовища та виділити три основні групи когнітивних викривлень: когнітивні, емоційно-афективні та

соціально-групові. Кожна група має характерні особливості прояву в умовах алгоритмічного контенту, інформаційного перевантаження та вірусності.

Таблиця 1.2.

Емоційно-афективні викривлення у контексті споживання контенту в кіберпросторі

№	Назва когнітивного викривлення	Психологічний механізм	Специфіка прояву в цифровому середовищі
1	Афективна евристика	Рішення та оцінки формуються переважно на основі емоційного ставлення, а не на об’єктивних фактах чи статистичних даних.	Яскравий, сенсаційний та короткий контент (меми, Reels, TikTok) максимально активізує Систему 1, роблячи емоційне забарвлення головним критерієм сприйняття інформації.
2	Евристика негативного упередження	Негативна інформація привертає значно більше уваги та швидше обробляється мозком порівняно з позитивною.	Негативний контент поширюється в соціальних мережах у 2–3 рази швидше, що активно використовується для маніпуляцій і провокування емоційного зараження.
3	Емоційне зараження	Емоції, закладені в повідомленні, швидко передаються між користувачами, викликаючи подібну реакцію в аудиторії.	Короткі емоційно насичені відео, меми та пости з провокативними образами забезпечують швидке поширення емоцій (страх, гнів, обурення) у цифровому середовищі.
4	Упередження оптимізму/песимізму	Емоційний стан індивіда суттєво спотворює оцінку ймовірності позитивних або негативних подій.	У періоди кризи емоційно заряджений контент посилює колективний песимізм або ілюзорний оптимізм, суттєво впливаючи на суспільні настрої та готовність до дій.

Таблиця 1.3.

Соціально-групові викривлення у контексті споживання контенту в кіберпросторі

№	Назва когнітивного викривлення	Психологічний механізм	Специфіка прояву в цифровому середовищі
1	Ефект приєднання до більшості	Схильність людини приймати думку або поведінку, яку демонструє більшість, через потребу в соціальному схваленні.	Візуальні індикатори соціального схвалення (лайки, репости, коментарі, кількість переглядів) суттєво посилюють ефект, створюючи ілюзію загальної думки та прискорюючи поширення наративів.
2	Упередження авторитету	Надмірна довіра до інформації, що походить від джерел, які сприймаються як авторитетні, незалежно від фактичної компетентності.	У цифровому середовищі «авторитетами» часто виступають інфлюенсери, популярні Telegram-канали та відомі акаунти, а не інституції, що полегшує маніпуляцію масовою свідомістю.
3	Групова поляризація	Посилення початкових позицій членів групи в бік більш радикальних поглядів після обговорення в групі.	Закриті спільноти, ехокамери та алгоритмічні стрічки (особливо в Telegram-каналах) сприяють радикалізації позицій і посиленню міжгрупової ворожості.
4	Соціальне порівняння	Оцінка інформації та власних поглядів відбувається через порівняння з «своїми» або «чужими» соціальними групами.	Цифрові платформи чітко маркують контент як «наш» чи «їхній», що посилює групову ідентифікацію та спотворює об’єктивне сприйняття інформації.

Запропонована класифікація за механізмами впливу дозволяє чітко диференціювати природу когнітивних викривлень у кіберпросторі та виявляти найбільш уразливі аспекти сприйняття інформації користувачами. Вона враховує специфіку цифрового середовища та створює основу для подальшого аналізу їхнього практичного прояву в умовах споживання різних типів контенту.

Таблиця 1.4.

Вплив алгоритмічної персоналізації на посилення когнітивних викривлень у кіберпросторі

№	Назва ефекту / механізму	Психологічний механізм	Специфіка прояву через алгоритмічну персоналізацію	Основні наслідки
1	Підтверджувальне упередження	Автоматична перевага інформації, що узгоджується з наявними переконаннями, та ігнорування суперечливих даних.	Рекомендаційні алгоритми (Facebook, Instagram, TikTok, YouTube, Telegram) постійно пропонують контент, що відповідає попереднім уподобанням користувача, пріоритезуючи матеріали з високою залученістю.	Формування «інформаційних бульбашок», звуження інформаційного горизонту, зниження критичного мислення
2	Ефект ехокамери	Посилення існуючих переконань через спілкування переважно з однодумцями.	Алгоритми створюють замкнене інформаційне середовище, в якому користувач отримує переважно контент від «однодумців», особливо в закритих Telegram-каналах та тематичних групах.	Радикалізація поглядів, сегментація суспільства, посилення поляризації
3	Ефект фільтраційної бульбашки	Ізоляція користувача в інформаційному просторі, що відповідає його інтересам і поглядам.	Персоналізовані стрічки значно обмежують доступ до альтернативних точок зору; користувач перестає навіть підозрювати про існування іншої інформації.	Зниження сприйняття альтернативних поглядів, зростання впевненості у власній правоті, фрагментація суспільної картини світу

Для наочного відображення впливу алгоритмічної персоналізації на посилення когнітивних викривлень у кіберпросторі розроблено схему, яка систематизує основні механізми цього впливу (рис. 1.2).

Рис. 1.2 Класифікація за домінуючим механізмом впливу на сприйняття інформації.

Механізми дії цих викривлень чітко проявляються при споживанні різних типів контенту:

– При споживанні новин алгоритми YouTube та Google News швидко фіксують тематичні інтереси та політичні уподобання, формуючи стрічку, де домінують матеріали одного спрямування.

– При перегляді відеоконтенту (TikTok, Reels, Shorts) персоналізація відбувається ще швидше через аналіз емоційної реакції (час перегляду, повторні перегляди), що особливо активно посилює confirmation bias та emotional contagion.

– При гортанні стрічок (Instagram, Facebook) алгоритми створюють безперервний потік контенту, який максимально відповідає інтересам, зменшуючи ймовірність зіткнення з дискомфортною або суперечливою інформацією.

Наслідки для індивідуального сприйняття є глибокими та багатоплановими.

По-перше, відбувається фрагментація картини світу — користувач втрачає здатність бачити складність соціальних явищ.

По-друге, посилюється поляризація поглядів, що ускладнює діалог між різними соціальними групами.

По-третє, формується когнітивна залежність від алгоритмів: людина поступово делегує функцію відбору інформації технологічним платформам, ще більше послаблюючи Систему 2.

В умовах гібридної війни такі ефекти мають стратегічне значення, оскільки дозволяють зовнішнім акторам цілеспрямовано впливати на окремі сегменти суспільства через точкову персоналізацію контенту. Виходячи з цього, когнітивні викривлення, посилені алгоритмічною персоналізацією, перетворюють звичайні психологічні механізми на потужний інструмент масового впливу. Їхнє розуміння є критично важливим для розробки ефективних заходів протидії в цифровому середовищі.

Характер прояву когнітивних викривлень значною мірою залежить від типу контенту, який споживає користувач. Різні формати (текстовий, візуальний, аудіовізуальний та інтерактивний) активізують різні психологічні механізми та посилюють специфічні спотворення. Це зумовлено особливостями сприйняття інформації мозком та архітектурою платформ.

Сучасний етап розвитку цифрових технологій створює принципово нові виклики для протидії соціально-психологічним наслідкам когнітивних викривлень. Специфіка прояву когнітивних викривлень залежно від типу цифрового контенту груповано у табл. 1.5

Таблиця 1.5.

Специфіка прояву когнітивних викривлень залежно від типу цифрового контенту

№	Тип контенту	Основні активовані когнітивні викривлення	Специфіка впливу на когнітивну систему	Основні наслідки
1	Текстовий контент (новинні статті, пости, аналітичні матеріали)	Підтверджувальне упередження, ефект якоря, евристика доступності	Користувач має час на осмислення, що сприяє вибірковій інтерпретації. Водночас через перевантаження переважає поверхневе читання (скімінг). Емоційно забарвлені заголовки швидко активізують упередження.	Посилення confirmation bias на платформах Facebook та Telegram; вибіркове сприйняття фактів.
2	Візуальний контент (меми, інфографіка, фотографії, графіка)	Евристика репрезентативності, афективна евристика, емоційне зараження, ефект бандвагона	Візуальні образи обробляються мозком значно швидше за текст і викликають миттєву емоційну реакцію.	Ефективне формування стійких уявлень про події; суттєве підвищення емоційного зараження та соціального доказу
3	Аудіовізуальний контент (відео, Reels, TikTok, YouTube)	Евристика доступності, емоційне зараження, афективна евристика	Поєднання зображення, звуку, музики та динаміки максимально залучає Систему 1. Короткі формати (15–60 секунд) майже не залишають часу для Системи 2.	Найвищий рівень впливу на когнітивну систему; синхронне емоційне та когнітивне навантаження
4	Інтерактивний контент (опитування, сториз, живі ефіри, коментарі)	ефект бандвагона, упередження авторитету, групова поляризація	Користувач безпосередньо взаємодіє з контентом і бачить реакції інших у реальному часі, що створює сильний соціальний тиск.	Максимальна активація соціально-групових викривлень; посилення відчуття колективної думки та соціального схвалення.

Найбільш загрозливими серед них є швидкий розвиток штучного інтелекту, діпфейк-технологій та алгоритмічного посилення когнітивних спотворень.

Штучний інтелект радикально підвищує ефективність маніпулятивного впливу. Генеративні моделі дозволяють створювати персоналізований контент, який ідеально адаптується під підтверджувальне упередження, емоційний стан та психологічний профіль конкретного користувача або соціальної групи. Це робить маніпуляцію масовою свідомістю більш точною та менш помітною.

Діпфейк-технології створюють додаткову загрозу, оскільки підривають базову довіру до візуальної інформації. Реалістичні фальшиві відео з відомими політиками, військовими та громадськими діячами здатні викликати потужне емоційне зараження та групову поляризацію за лічені години. Вони ирадикально посилюють ефект авторитету та емоційне зараження, оскільки фальшиве відео з відомою особою сприймається як достовірне на рівні автоматичного сприйняття. Розвиток ШІ перетворює когнітивні викривлення на високоточну зброю в інформаційній війні, оскільки фальшивий контент стає практично невідрізним від реального для більшості користувачів.

Специфіка прояву когнітивних викривлень суттєво відрізняється залежно від типу контенту та платформи. Найнебезпечнішим є поєднання аудіовізуального та інтерактивного форматів з технологіями ШІ, що створює принципово нові виклики для когнітивної безпеки користувачів.

Відмінності поведінки користувачів на різних платформах:

- TikTok — максимальна активація емоційно-афективних викривлень через алгоритм, орієнтований на утримання уваги. Користувачі демонструють найвищий рівень імпульсивного споживання.
- Telegram — поєднання текстового та аудіовізуального контенту в закритих каналах сприяє посиленню ехокамер та групової поляризації
- Facebook — старша аудиторія, довші тексти та обговорення, що активізує confirmation bias і authority bias.

- Twitter/X — швидкий темп, короткі повідомлення та тренди посилюють евристику доступності та ефект бандвагона

Підсумовуюч, когнітивні викривлення в умовах сучасного кіберпростору набувають нових, специфічних форм, які потребують окремої систематизації. .

Значення розробленої класифікації полягає в тому, що вона дозволяє перейти від загального розуміння природи когнітивних викривлень до конкретного аналізу їхньої дії в реальному цифровому середовищі. Вона розкриває психологічні механізми вразливості сучасної людини перед інформаційними загрозами, показуючи, як різні типи контенту та платформи по-різному експлуатують психологічні особливості сприйняття. Такий підхід є особливо актуальним для України в умовах триваючої гібридної війни, де когнітивні викривлення стають інструментом масового впливу на суспільну свідомість.

1.3. Соціально-психологічний вплив когнітивних викривлень на суспільство

У сучасних умовах цифрової трансформації суспільства когнітивні викривлення переходять із сфери індивідуальної психології на макрорівень, стаючи потужним фактором впливу на соціальні групи, громадську думку, політичні процеси та загальну стабільність суспільства. Якщо на індивідуальному рівні когнітивні спотворення вивчені досить детально, то їхні соціально-психологічні наслідки на рівні суспільства залишаються недостатньо систематизованими. Саме тому дослідження впливу когнітивних викривлень на макрорівні набуває особливої актуальності.

Масовий характер впливу когнітивних спотворень у кіберпросторі призводить до масштабних соціальних наслідків. За даними досліджень, у країнах з високим рівнем проникнення соціальних мереж систематичне посилення підтверджувального упередження та групової поляризації призводить до зростання суспільної поляризації, зниження довіри до державних інститутів

та зростання вразливості до маніпулятивних інформаційних кампаній. Ecker K.H. зазначає, що стійкість до корекції дезінформації на індивідуальному рівні суттєво ускладнює подолання колективних хибних уявлень, які закріплюються в суспільній свідомості [15].

Особливої гостроти проблема набуває в умовах гібридної війни. Російська Федерація активно використовує когнітивні викривлення як інструмент інформаційно-психологічних операцій, спрямованих на розкол українського суспільства, підрив довіри до влади, армії та західних партнерів.

Статистичні дані підтверджують критичний масштаб проблеми. Згідно з дослідженнями 2024–2025 років, понад 65% українців регулярно стикаються з маніпулятивним контентом, який експлуатує когнітивні викривлення, що призводить до зростання суспільної тривожності, поляризації та зниження соціальної згуртованості [16].

Яскравим підтвердженням критичного масштабу проблеми є динаміка рівня медіаграмотності українського суспільства. Незважаючи на певні позитивні зрушення, значна частка населення продовжує демонструвати низький або середній рівень критичного сприйняття інформації, що суттєво підвищує вразливість до маніпулятивного контенту (рис. 1.3).

У 2022 році спостерігалося суттєве зростання рівня медіаграмотності українців (частка осіб з високим та вищим за середній рівнем зросла до 81%). Однак уже у 2023–2024 роках відбулося певне зниження показників. У 2024 році лише 7% населення мали високий рівень медіаграмотності, а частка осіб з низьким та нижчим за середній рівнем зросла до 28%. Середній індекс медіаграмотності, який у 2022 році сягнув піку 5,8 бала, у 2024 році знизився до 5,2 бала.

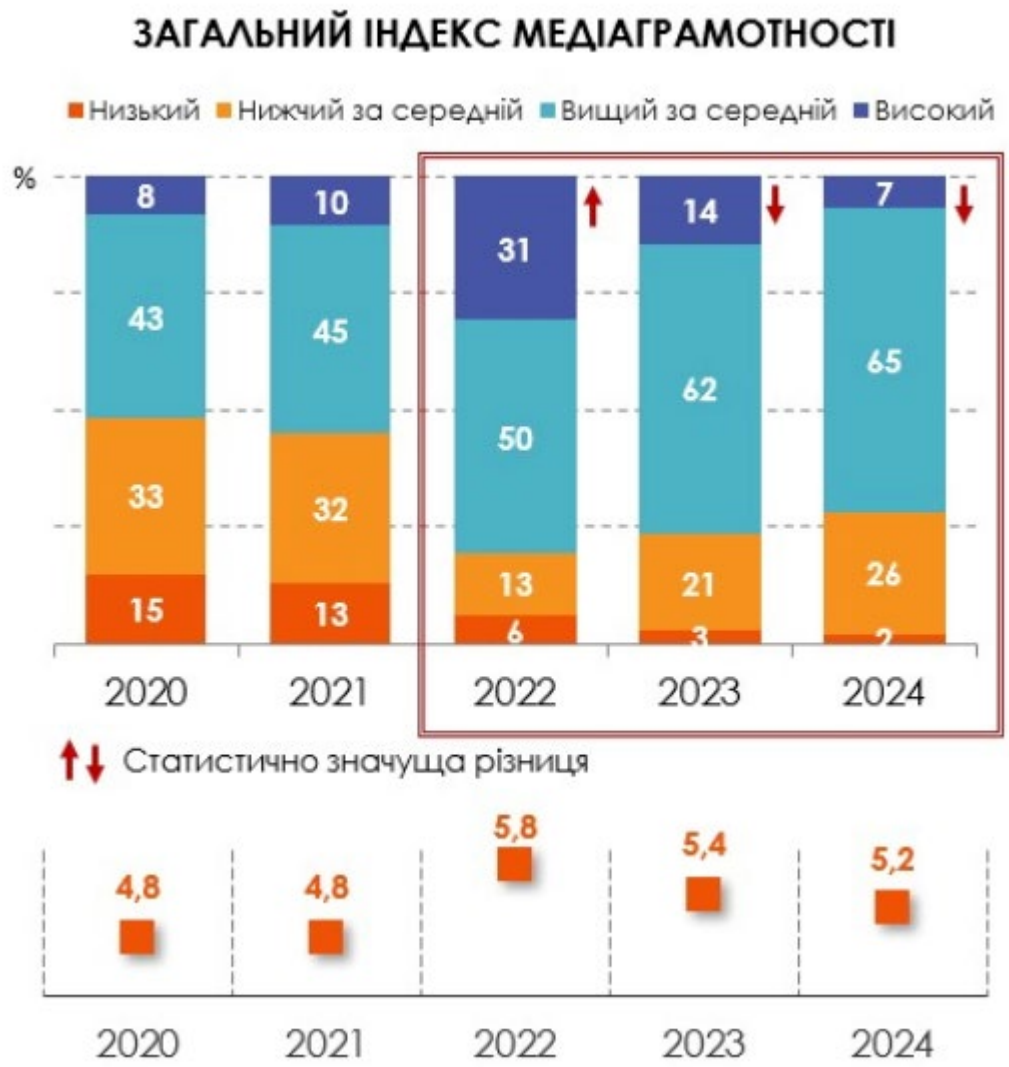

Рис. 1.3. Динаміка загального індексу медіаграмотності населення України у 2020–2024 роках

Як свідчать дані з рис. 1.1, існує виразна зворотна залежність між рівнем медіаграмотності населення та його вразливістю до маніпулятивного контенту. Зниження індексу медіаграмотності у 2023–2024 роках призвело до зростання частки населення, яке піддається впливу когнітивних викривлень

Виходячи з цього можна константувати, що когнітивні викривлення суттєво впливають на розвиток інформаційних відносин і економічну поведінку населення, що в умовах війни може мати руйнівні наслідки для економічної стійкості країни.

Перехід від індивідуальних когнітивних спотворень до колективних соціально-психологічних ефектів є одним із найбільш складних і водночас ключових аспектів вивчення впливу когнітивних викривлень на суспільство.

Теоретичні основи такого переходу ґрунтуються на синтезі когнітивної психології та соціальної психології. Індивідуальні когнітивні спотворення, будучи універсальними для більшості людей, за певних умов починають синхронізуватися в соціальних групах і масштабуватися на рівні суспільства. Цей процес відбувається через механізми соціальної взаємодії, які перетворюють особисті упередження на спільні колективні уявлення.

Таблиця 1.6.

Основні механізми переходу індивідуальних когнітивних викривлень у колективні соціально-психологічні ефекти

№	**Назва механізму**	**Психологічний механізм**	**Специфіка прояву в цифровому середовищі та умовах гібридної війни**	**Ключові наслідки**
1	Емоційне зараження	Швидка передача емоцій (страх, гнів, обурення) між індивідами, що випереджає раціональний аналіз.	Цифрові платформи забезпечують миттєве поширення емоційно заряджених повідомлень, мемів та коротких відео, перетворюючи індивідуальні емоційні реакції на масові психологічні стани.	Каталізатор масового поширення когнітивних спотворень; формування колективної тривожності або агресії
2	Соціальний доказ	Схильність індивідів сприймати поведінку або думку більшості як правильну, особливо в умовах невизначеності.	Візуальні індикатори (лайки, репости, коментарі, кількість переглядів) суттєво посилюють ефект, сприяючи поширенню наративів через ефект бандвагона та групової поляризації	Посилення колективних упереджень; маніпуляція масовою свідомістю в умовах когнітивної війни
3	Колективна ірраціональність	Масштабування суми індивідуальних когнітивних спотворень у стійкі колективні хибні уявлення.	Алгоритмічні ехокамери та соціальне підкріплення роблять колективні помилкові переконання значно стійкішими до спростування, ніж індивідуальні.	Формування стійких соціальних патологій; ускладнення корекції дезінформації через групову ідентичність

Теоретичні основи соціально-психологічного впливу когнітивних викривлень базуються на розумінні механізмів масштабування індивідуальних спотворень через емоційне зараження, соціальний доказ та групову поляризацію Вони відіграють визначальну роль у процесах формування громадської думки та політичних настроїв у сучасному суспільстві. У цифровому середовищі індивідуальні психологічні механізми масштабуються, перетворюючись на потужні фактори колективного сприйняття реальності. Найбільш впливовими серед них є підтверджувальне упередження, ефект ехокамер та групова поляризація, які спільно сприяють поляризації суспільства, ерозії довіри до державних інститутів та поширенню маніпулятивних наративів.

Когнітивні викривлення в інформаційній війні сприяють штучному розколу суспільства за політичними, регіональними або ціннісними ознаками. У сучасних умовах це проявляється у глибокому розподілі українського суспільства щодо питань зовнішньої політики, ставлення до мобілізації, оцінки діяльності влади тощо. Поляризація знижує здатність суспільства до консолідації перед зовнішніми загрозами.

Значним наслідком є ерозія довіри до державних інститутів. Постійне підживлення підтверджувального упередження через маніпулятивні наративи призводить до системного недовіри до ЗМІ, влади, армії та міжнародних партнерів. Окремо зазначається, що емоційні тригери та когнітивні спотворення в інформаційному просторі України суттєво підривають довіру населення до офіційних джерел інформації, замінюючи їх на альтернативні, часто маніпулятивні канали.

Статистичні дані свідчать про серйозність ситуації. За результатами соціологічних досліджень 2024–2025 років, рівень поляризації українського суспільства за ключовими суспільно-політичними питаннями зріс на 28–35% порівняно з довоєнним періодом [17]. Близько 60% респондентів відзначають, що вони уникають обговорення політичних тем з людьми, які мають інші погляди

У сучасній гібридній війні когнітивні викривлення перетворилися на один із найефективніших інструментів інформаційно-психологічних операцій (ІПСО). На відміну від традиційних військових засобів, вони дозволяють впливати на свідомість населення без прямого фізичного насильства, досягаючи стратегічних цілей через масову маніпуляцію сприйняттям реальності.

Моральний стан суспільства також зазнає значного негативного впливу. Постійне використання емоційного зараження через страх, безнадійність, зраду та втомлення призводить до зростання апатії, цинізму та деморалізації. Емоційні тригери в інформаційному просторі України провокують колективні психологічні стани, що знижують мотивацію до опору та підтримки ЗСУ. Особливо небезпечним є ефект «нормалізації агресора» — поступове прийняття російських наративів частиною населення через евристику доступності та підтверджувальне упередження.

Готовність до опору також зазнає впливу. З одного боку, успішне використання когнітивних викривлень може тимчасово підвищувати мотивацію через ефект «об'єднання перед зовнішньою загрозою». З іншого — системне застосування деморалізуючих наративів призводить до «втоми від війни», зниження віри у перемогу та готовності терпіти труднощі.

Масовий вплив когнітивних викривлень виходить далеко за межі політичної та воєнної сфери та призводить до глибоких системних наслідків для економіки, культури, психологічного здоров'я населення та загальної суспільної стійкості. Системне поширення спотворень через цифрові канали трансформує індивідуальні упередження в колективні соціальні патології, що негативно позначається на різних сферах суспільного життя.

Когнітивні спотворення суттєво впливають на розвиток ринкових відносин, спотворюючи економічну поведінку як окремих громадян, так і бізнесу. Підтверджувальне упередження та евристика доступності призводять до ірраціональних інвестиційних рішень, панічних закупівель, бойкотів окремих брендів чи, навпаки, сліпої довіри до сумнівних фінансових пропозицій. У період війни це проявляється в поширенні фінансових шахрайств, недовірі до

національної валюти, панічних відтоках капіталу та зниженні інвестиційної привабливості країни. Масова віра в маніпулятивні наративи про «економічний крах» або «неминучу поразку» може призводити до реального погіршення економічних показників через самовиконувані прогнози.

Культурні наслідки пов'язані з руйнуванням спільної культурної ідентичності та ціннісного простору. Постійне функціонування ехокамер і групової поляризації сприяє фрагментації культурного поля: різні соціальні групи формують власні, часто взаємовиключні, культурні наративи. Масове поширення когнітивних викривлень призводить до «культурної балканізації» суспільства, коли спільна національна ідентичність поступово розмивається під тиском альтернативних, часто зовнішньо інспірованих, культурних кодів. В українському контексті це проявляється у протистоянні щодо мови, історичної пам'яті, ставлення до традицій та сучасної культури, що послаблює культурну єдність нації.

Міжпоколінні відносини також зазнають значної деформації. Молодше покоління, більш активно залучене в алгоритмічні платформи, частіше потрапляє під вплив підтверджувального упередження та емоційного зараження, формуючи світогляд, який суттєво відрізняється від світогляду старших поколінь. Це призводить до зростання взаємного непорозуміння, конфліктів цінностей та зниження міжпоколінної солідарності — важливого фактора суспільної стійкості в умовах війни.

Системні когнітивні викривлення активно сприяють соціальній фрагментації та зниженню суспільної стійкості. Коли різні соціальні групи сприймають реальність крізь призму різних, часто протилежних, інформаційних бульбашок, здатність суспільства до консолідованої дії суттєво зменшується. *Т*ривала дія когнітивних спотворень призводить до послаблення довіри, взаємодопомоги та готовності до спільних дій. У результаті суспільство стає більш вразливим до зовнішніх і внутрішніх потрясінь.

Таким чином, економічні, культурні та психологічні наслідки масового впливу когнітивних спотворень мають комплексний і взаємопов'язаний

характер. Вони не лише знижують економічну ефективність, культурну єдність та психологічне здоров'я населення, але й підривають фундаментальні основи суспільної стійкості.

Ефективна протидія когнітивним загрозам у сучасних умовах гібридної війни можлива лише за умови системного поєднання індивідуального та колективного рівнів захисту. Оскільки просте реактивне спростування дезінформації часто виявляється недостатнім і навіть може спричиняти ефект бумеранга, пріоритетним напрямком стає впровадження проактивних стратегій психологічної імунізації (inoculation) населення. У відповідь на ці виклики необхідним є комплексний підхід до підвищення суспільної когнітивної стійкості, який передбачає інтеграцію освітніх, технічних, інституційних та психологічних заходів (табл 1.7)

Таблиця 1.7.

Комплексний підхід до підвищення суспільної когнітивної стійкості

№	Компонент підходу	Основний зміст та завдання	Рівень реалізації	Очікуваний ефект
1	Освітні заходи	Системне впровадження програм розвитку критичного мислення, медіаграмотності та розуміння механізмів когнітивних спотворень на всіх рівнях освіти.	Індивідуальний / Груповий	Формування довгострокової когнітивної стійкості населення, особливо серед молоді.
2	Технічні рішення	Розробка та впровадження інструментів детекції deepfake, маніпулятивного контенту та алгоритмічного моніторингу інформаційного простору.	Інституційний / Технічний	Оперативне виявлення загроз і зменшення швидкості поширення маніпулятивного контенту.
3	Інституційні механізми	Створення державних стратегій когнітивної безпеки, координація зусиль державних органів, громадських організацій та технологічних компаній.	Інституційний	Системна державна політика захисту інформаційного простору та національної стійкості.
4	Психологічні програми	Розробка та реалізація програм підвищення емоційної стійкості населення та протидії емоційному зараженню.	Індивідуальний / Груповий	Зниження вразливості до емоційних тригерів і підвищення психологічної стійкості суспільства.

Такий багаторівневий підхід дозволяє не лише реагувати на вже поширені маніпуляції, але й формувати стійкість до майбутніх інформаційних загроз на рівні окремої людини, соціальних груп та суспільства в цілому.

Сучасні виклики, пов’язані зі стрімким розвитком штучного інтелекту, генеративних моделей та алгоритмічних технологій персоналізації контенту, радикально змінюють характер інформаційних загроз і об’єктивно вимагають переходу від переважно реактивних заходів протидії (таких як фактчекінг і спростування вже поширеного маніпулятивного контенту) до проактивної, системної стратегії підвищення когнітивної стійкості суспільства.

Якщо раніше основним завданням було виявлення та спростування фейкової інформації, то сьогодні, коли ШІ дозволяє створювати високоякісний, персоналізований та емоційно адаптований контент у реальному часі, просте реагування на вже опубліковані матеріали стає недостатнім.

Саме тому постає нагальна необхідність формування національного психологічно-превентивного інструменту населення через комплекс освітніх, психологічних та інституційних заходів, спрямованих на підвищення критичного мислення, емоційної регуляції та стійкості до маніпулятивного впливу ще до моменту зіткнення з ним. Такий проактивний підхід дає змогу не лише зменшити ефективність ворожих інформаційно-психологічних операцій, але й зміцнити загальну суспільну стійкість у довгостроковій перспективі.

Висновки до розділу 1

У цьому розділі проведено теоретичний аналіз основ когнітивних викривлень та особливостей їх прояву в умовах сучасного кіберпростору. Було розглянуто концепцію систем мислення, природу когнітивних спотворень, їх класифікацію в цифровому середовищі та соціально-психологічний вплив на суспільство.

Було досліджено двоїсту модель мислення Д. Канемана (Система 1 і Система 2), природу евристик та когнітивних спотворень. Показано, що

когнітивні викривлення є не випадковими помилками, а системними особливостями людського мислення, сформованими в процесі еволюції. В умовах кіберпростору ці механізми суттєво посилюються через інформаційне перевантаження, алгоритмічну персоналізацію та емоційне зараження

Також розроблено багатофакторну класифікацію когнітивних викривлень у контексті споживання контенту в кіберпросторі. Виділено три основні групи: когнітивні, емоційно-афективні та соціально-групові викривлення. Проаналізовано специфіку їх прояву залежно від типу контенту (текстовий, візуальний, аудіовізуальний, інтерактивний) та вплив алгоритмічної персоналізації (інформаційні бульбашки, ехокамери, фільтраційні бульбашки).

Особливу увагу приділено соціально-психологічному впливу когнітивних викривлень на суспільство. Встановлено наявність зворотної залежності між рівнем медіаграмотності населення та його вразливістю до маніпулятивного контенту. Зниження індексу медіаграмотності у 2023–2024 роках супроводжувалося зростанням частки населення, схильного до впливу когнітивних спотворень, що особливо небезпечно в умовах триваючої гібридної війни.

Теоретичний аналіз показав, що когнітивні викривлення є одним із ключових факторів вразливості людини та суспільства в цифровому середовищі. Глибоке розуміння їхньої природи, механізмів прояву та соціально-психологічних наслідків створює необхідне методологічне підґрунтя для подальшого дослідження технік маніпулювання та розробки ефективних методів протидії.

Розділ 2 АНАЛІЗ МЕХАНІЗМІВ ВИКОРИСТАННЯ КОГНІТИВНИХ ВИКРИВЛЕНЬ В ІНФОРМАЦІЙНИХ ОПЕРАЦІЯХ

2.1. Когнітивні викривлення як інструмент реалізації ІПСО

Сучасна геополітична ситуація характеризується переходом від традиційних форм збройного протистояння до гібридних війн, у яких інформаційно-психологічні операції (ІПСО) відіграють одну з провідних ролей. У цих умовах когнітивні викривлення перетворилися на потужний і ефективний інструмент досягнення стратегічних цілей без необхідності прямого військового зіткнення. Актуальність вивчення цієї проблеми зумовлена стрімким зростанням впливу інформаційного компонента у сучасних конфліктах, зокрема у російсько-українській війні.

Гібридна війна Росії проти України, що триває з 2014 року та перейшла у повномасштабну фазу у 2022 році, яскраво демонструє, як когнітивні викривлення стають ключовим елементом військової стратегії. В сучасній когнітивній війні людська свідомість стала головним полем бою, а маніпуляція когнітивними процесами — одним із найбільш ефективних способів впливу на противника. На відміну від класичних військових дій, ІПСО, що ґрунтуються на експлуатації когнітивних викривлень, дозволяють досягати значних результатів із мінімальними матеріальними витратами.

За даними досліджень, роль інформаційно-психологічних операцій у сучасних конфліктах постійно зростає. Когнітивні операції стали невід’ємною частиною сучасної гібридної війни, демонструючи високу ефективність у формуванні громадської думки, деморалізації противника та розколі суспільства. У звіті «Cognitive Attacks in Russian Hybrid Warfare» (2018) підкреслюється, що Росія активно застосовує когнітивні атаки як системний елемент своєї стратегії, спрямований на підрив внутрішньої стабільності країн-суперників [18].

Когнітивні викривлення стали ключовим інструментом досягнення стратегічних цілей без прямого військового зіткнення з кількох причин. По-

перше, вони дозволяють впливати на рішення мільйонів людей одночасно, формуючи потрібну інтерпретацію подій. По-друге, їхній вплив важко відстежити та спростувати у реальному часі. По-третє, вони мають пролонгований ефект, оскільки закріплюються на рівні колективної свідомості. Bârgăoanu A. (2024) підкреслює, що сучасні когнітивні операції спрямовані не стільки на зміну конкретних поглядів, скільки на створення стійких когнітивних структур, які продовжують працювати навіть після завершення активної фази операції [19].

Трансформація когнітивних спотворень із психологічного феномена в інструмент інформаційно-психологічного впливу є одним із ключових процесів сучасної гібридної війни. Якщо в класичній психології когнітивні викривлення розглядалися переважно як індивідуальні особливості мислення, то в контексті інформаційних операцій вони набувають характеру цілеспрямовано експлуатованого ресурсу масового впливу.

Теоретичні засади такого застосування ґрунтуються на синтезі когнітивної психології, теорії комунікації та сучасних концепціях гібридної війни. Сучасні дослідження визначають когнітивну війну як протистояння, де головним об’єктом ураження є свідомість противника, а основною зброєю — системна маніпуляція когнітивними процесами. На відміну від традиційних форм війни, когнітивна війна спрямована не на фізичне знищення, а на зміну сприйняття реальності, цінностей та мотивації цільової аудиторії.

Однією з фундаментальних теорій у цій сфері є концепція когнітивних операцій у гібридній війні. Danyk Y. та Briggs Ch.M. (2023) у роботі «Modern Cognitive Operations and Hybrid Warfare» розглядають когнітивні операції як інтегрований елемент гібридної стратегії, де інформаційний вплив доповнює або навіть замінює кінетичні дії [20]. Автори підкреслюють, що ефективність таких операцій залежить від точного використання психологічних уразливостей — насамперед когнітивних викривлень — для формування бажаного сприйняття подій.

Дослідження показують, що ІПСО будуються на системному використанні евристик і упереджень для створення «когнітивних пасток», які змушують цільову аудиторію самостійно доходити до потрібних агресору висновків. Такий підхід є більш ефективним, ніж пряма пропаганда, оскільки зменшує критичне сприйняття інформації.

Механізм переходу від індивідуальних когнітивних упереджень до масового маніпулятивного впливу відбувається у кілька етапів:

1. Ідентифікація вразливостей — визначення домінуючих когнітивних спотворень у цільових соціальних групах (підтверджувальне упередження, страх втрат, евристика доступності тощо).

2. Адаптація контенту — створення повідомлень, які максимально відповідають психологічним особливостям аудиторії.

3. Масштабування через цифрові платформи — використання алгоритмів для поширення контенту в ехокамерах і інформаційних бульбашках.

4. Закріплення ефекту — повторення наративів для формування стійких колективних переконань.

Такий механізм дозволяє перетворювати природні психологічні особливості людини на інструмент зовнішнього управління поведінкою.

Особливу роль когнітивні викривлення відіграють як елемент «м'якої сили» у сучасних інформаційних операціях. На відміну від «жорсткої сили» (військової чи економічної), «м'яка сила» діє через переконання та формування бажаних інтерпретацій. Когнітивні викривлення у цьому контексті виступають як каталізатор, що робить вплив більш тонким і менш помітним. Успішне використання когнітивних механізмів дозволяє досягати ефекту, подібного до військової перемоги, без значних матеріальних втрат і міжнародного осуду.

Ефективність інформаційно-психологічних операцій (ІПСО) значною мірою залежить від системного та цілеспрямованого використання когнітивних викривлень. Сучасні ІПСО не обмежуються простим поширенням дезінформації, а будуються на глибокому розумінні психологічних механізмів і їхній експлуатації на різних етапах операції.

Будь-яка ефективна когнітивна операція проходить три основні етапи:

1. Підготовка — аналіз цільової аудиторії, ідентифікація домінуючих когнітивних уразливостей, створення контенту, адаптованого під конкретні спотворення.

2. Поширення — масове впровадження контенту через цифрові канали з урахуванням алгоритмів платформ.

3. Закріплення — повторення та посилення наративів для переходу від короткострокового впливу до стійких когнітивних змін.

Російські інформаційно-психологічні операції проти України є одним із найбільш системних і довготривалих прикладів використання когнітивних викривлень як інструменту гібридної війни (табл.2.1)

Таблиця 2.1.

Еволюція основних наративів російських ІПСО проти України (2014–2025 рр.)

Період	**Етап**	**Основні наративи**	**Експлуатовані когнітивні викривлення**	**Характеристика та цілі**
2014–2021	Підготов-чий	«Україна — штучна держава», «українці та росіяни — один народ», «на Донбасі — громадянська війна»	Підтверджувальне упередження, евристика репрезентативності, соціальне порівняння	Формування спотвореної історичної та ідентичної картини, створення ґрунту для подальшої дестабілізації
2022	Повномас-штабне вторгнення	«Київ за три дні», «ЗСУ розгромлені», «спецоперація з денацифікації», «російська армія визволитель»	Емоційне зараження (страх, паніка), евристика доступності, упередження оптимізму (для російської аудиторії)	Швидке деморалізування українського суспільства, створення ефекту неминучої поразки, виправдання агресії
2023–2025	Затяжна війна	«Захід зрадив Україну», «влада кидає народ на безнадійні штурми», «корупція в армії», «неминучість поразки», «Україна — провальний проєкт»	Підтверджувальне упередження, групова поляризація, негативне упередження, упередження безнадійності (Learned Helplessness)	Підтримання деморалізації, ерозія довіри до влади та партнерів, посилення внутрішньої поляризації суспільства

На відміну від класичної пропаганди, російські ІПСО будуються на глибокому розумінні психологічних уразливостей українського суспільства та їхній цілеспрямованій експлуатації. Головною метою є не стільки поширення брехні, скільки зміна способу мислення та сприйняття реальності значної частини населення. Сегментація цільових аудиторій представлена у табл. 2.2.

Таблиця 2.2.

Сегментація цільових аудиторій російських ІПСО та адаптація когнітивних викривлень

Цільова аудиторія	**Психологічні та соціальні особливості**	**Основні експлуатовані когнітивні викривлення**	**Характерні наративи та приклади**	**Мета впливу**
Населення Сходу та Півдня України	Регіональна ідентичність, ностальгія за СРСР, економічні проблеми, мовні особливості	Підтверджувальне упередження, соціальне порівняння, евристика репрезентативності	«Тут завжди говорили російською», «Україна — це гноблення Донбасу», «Ми — один народ»	Посилення сепаратистських настроїв, створення внутрішнього розколу
Молодь (14–25 років)	Висока цифрова активність, короткий цикл уваги, схильність до емоційного контенту	Евристика доступності, емоційне зараження, афективна евристика	Короткі TikTok/Reels-відео з гумористичним або шокуючим контентом про «безглуздість війни», меми про мобілізацію	Деморалізація, зниження мотивації до опору, формування антивоєнних настроїв
Жінки та матері	Висока емоційна чутливість, страх за близьких, материнський інстинкт	Негативне упередження, емоційне зараження, упередження безнадійності	Наративи «м’ясні штурми», «матері ховають синів через владу», історії «вдів і сиріт»	Підрив морально-психологічного стану суспільства, тиск на мобілізацію
Військовослужбовці та їхні родини	Високий рівень стресу, недовіра до командування в умовах затяжної війни	Упередження авторитету (у перевернутій формі), підтверджувальне упередження, ефект якоря	«Командування кидає на м’ясо», «генерали зраджують», «влада збагачується на крові»	Зниження дисципліни, деморалізація армії, поширення зради у військовому середовищі
Інтелектуальна та бізнес-еліта	Раціоналізація, прагматизм, страх економічних втрат	Упередження оптимізму/раціоналізації, підтверджувальне упередження, ефект якоря	«Опір безперспективний», «неминучість компромісу», «Україна — провальний проєкт Заходу»	Формування капітулянтських настроїв, послаблення елітної підтримки опору

Російські інформаційно-психологічні операції вирізняються високим рівнем сегментації цільових аудиторій та точної адаптації когнітивних викривлень під психологічні особливості, цінності та вразливості кожної групи. Такий диференційований підхід дозволяє максимально ефективно експлуатувати існуючі соціально-психологічні розломи українського суспільства. Нижче наведено основні цільові аудиторії та специфіку використання когнітивних викривлень щодо них.

Ефективність російських інформаційно-психологічних операцій не є сталою величиною і суттєво відрізняється залежно від етапу війни, суспільних настроїв, адаптивності українських механізмів протидії та зовнішніх факторів. Російські операції демонструють високу здатність до еволюції, проте їхній загальний стратегічний успіх залишається обмеженим. Для системного аналізу ефективності доцільно розглянути динаміку впливу російських ІПСО на різних етапах російсько-української війни.

Загальна ефективність російських когнітивних операцій може бути оцінена як середня з чіткою тенденцією до зниження. Незважаючи на значні ресурси, вкладені в інформаційно-психологічні операції, Росії не вдалося досягти головної стратегічної мети — масової деморалізації українського суспільства та примусу до капітуляції.

Водночас російським ІПСО вдалося підтримувати постійну внутрішню напругу, частково знижувати рівень суспільної довіри до влади, ускладнювати процеси мобілізації та створювати перешкоди для міжнародної підтримки України. Російські операції вирізняються високою адаптивністю — вони швидко переходять від грубих пропагандистських технік до більш витончених форм маніпуляції, зокрема часткового визнання проблем з одночасним перекладанням відповідальності. Ефективність російських ІПСО проти України за етапами війни наведена у табл. 2.3

Таблиця 2.3.

Ефективність російських ІПСО проти України за етапами війни

Період	**Рівень ефективності**	**Ключові прийоми російських ІПСО**	**Обмеження та прояви стійкості українського суспільства**	**Основні фактори впливу**
Лютий–березень 2022 (початок повномасштабного вторгнення)	Висока	Масова паніка, значна міграція населення, тимчасові проблеми з мобілізацією, хаос в інформаційному просторі	Швидке відновлення суспільної згуртованості після перших тижнів шоку	Ефект несподіванки, перевага в швидкості поширення інформації, використання емоційного зараження
2023–2024	Середня	Поширення наративів про корупцію, «м'ясні штурми», втому від війни; часткове зниження довіри до влади та партнерів	Відсутність стратегічного розколу суспільства, збереження високого рівня підтримки ЗСУ та опору	Адаптація українського суспільства, розвиток медіаграмотності, контрнаративи
2025 рік	Середня з тенденцією до зниження	Підтримання внутрішньої напруги, ускладнення мобілізаційних процесів, ерозія довіри до державних інститутів	Зростання стійкості до маніпуляцій, зниження сприйнятливості до грубих фейків	Втома від війни, економічні труднощі, еволюція українських механізмів протидії

Таким чином, особливості застосування когнітивних викривлень Росією проти України демонструють високу адаптивність і технологічність сучасних ІПСО. Розуміння цих особливостей є критично важливим для розробки ефективних контрзаходів, що буде розглянуто в подальших підрозділах роботи.

2.2 Упередження та маніпуляції в інформаційних кампаніях гібридної війни

Сучасна гібридна війна Росії проти України демонструє, що інформаційні кампанії стали не допоміжним, а одним із головних інструментів досягнення стратегічних цілей.

В умовах гібридної війни когнітивні операції досягли рівня стратегічного інструменту, порівнянного за впливовістю з традиційними військовими засобами. Згідно з аналізом NATO Defense College Foundation (2025), інформаційні кампанії, побудовані на експлуатації когнітивних упереджень, демонструють високу ефективність у підриві суспільної стійкості та створенні внутрішніх розколів у країнах-жертвах агресії [21]. У випадку України російські інформаційні операції системно використовують упередження для досягнення цілей, які раніше вимагали масштабних військових зусиль.

Відмінність сучасних маніпуляцій від класичної пропаганди є принциповою. Якщо традиційна пропаганда XX століття базувалася переважно на прямому нав'язуванні ідеологічних конструктів і цензурі, то сучасні маніпуляції в гібридній війні спрямовані на експлуатацію існуючих когнітивних упереджень людини. Вони не стільки нав'язують нову думку, скільки спрямовують мислення аудиторії у потрібне русло, використовуючи різні руйнівні психологічні методи

Сучасні теоретичні концепції поєднують досягнення когнітивної психології з теоріями інформаційної війни та рефлексивного управління. Найважливішими серед них є три ключові моделі:

Сучасні інформаційно-психологічні операції ґрунтуються на кількох ключових теоретичних моделях, які пояснюють механізми системного впливу на людську свідомість. Найбільш релевантними для аналізу російських ІПСО є три взаємодоповнювальні моделі: Cognitive Warfare, Reflexive Control та Hybrid Influence Operations.

1. Модель Cognitive Warfare (Когнітивна війна)

Концепція когнітивної війни розглядає людський розум як повноцінний домен воєнних дій, рівноправний із землею, морем, повітрям, космосом і кіберпростором [22].

Визначають когнітивну війну як системне застосування досягнень психології та нейронаук для цілеспрямованого впливу на процеси сприйняття, формування суджень і прийняття рішень супротивника. У межах цієї моделі когнітивні викривлення перестають бути лише індивідуальними психологічними вадами, а перетворюються на стратегічні «точки входу» (entry points), через які здійснюється масовий і прицільний вплив на свідомість.

2. Модель Reflexive Control (Рефлексивне управління)

Рефлексивне управління — це класична російська військово-наукова концепція, яка передбачає створення такого інформаційного середовища, за якого противник самостійно, керуючись власними логікою та упередженнями, приймає рішення, вигідні маніпулятору [23].

Рефлексивне управління особливо ефективне саме завдяки експлуатації когнітивних викривлень. Агресор не нав’язує готові висновки, а формує умови, в яких жертва, спираючись на підтверджувальне упередження, евристику доступності чи ефект якоря, обирає «свій» варіант поведінки, який насправді запрограмований супротивником.

3. Модель Hybrid Influence Operations (Гібридного операційного управління)

Ця модель акцентує увагу на комплексному, багатодоменному характері сучасного впливу, що поєднує інформаційні, психологічні, економічні, дипломатичні та кібернетичні інструменти.

Звіти NATO Defense College Foundation (2025) та Institute for Defence Studies and Analyses (2025) розглядають hybrid influence operations як інтегровані кампанії, в яких маніпулятивні техніки спеціально розробляються під виявлені когнітивні уразливості конкретних соціальних груп. Когнітивні викривлення в цій моделі слугують синергією різних інструментів впливу. Порівняльна характеристика теоретичних моделей інформаційного впливу у табл. 2.4

Таблиця 2.4.

Порівняльна характеристика ключових теоретичних моделей інформаційного впливу

Модель	**Основний об'єкт впливу**	**Роль когнітивних упереджень**	**Ключовий механізм маніпуляції**
Когнітивна війна	Свідомість та процеси мислення	Центральний елемент (точки уразливості)	Системна експлуатація психологічних механізмів
Рефлексивне управління	Прийняття рішень	Інструмент керування вибором жертви	Створення «когнітивних пасток»
Гібридне операційне управління	Суспільство в цілому	Підґрунтя для комплексного впливу	Поєднання різних рівнів маніпуляції

Основні типи взаємодії когнітивних упереджень і маніпулятивних технік наведено у табл. 2.5.

Таблиця 2.5.

Основні типи взаємодії когнітивних упереджень і маніпулятивних технік

Тип упередження	**Маніпулятивна техніка**	**Стратегічна мета в гібридній війні**	**Приклад застосування**
Підтверджувальне упередження	Ехокамери + персоналізований контент	Поляризація суспільства	Наративи «Зрада Заходу»
Евристика доступності	Вірусний контент + повторення	Формування спотвореної картини реальності	Перебільшення втрат ЗСУ
Емоційне зараження	Сенсаційні меми та короткі відео	Деморалізація та створення паніки	Кампанії «втома від війни»
Групова поляризація	Рефлексивне управління + соціальний доказ	Розкол суспільства	Протиставлення «влада vs народ»

Ефективність маніпулятивного впливу в інформаційних операціях значно зростає, коли він ґрунтується не на прямій фальсифікації фактів, а на цілеспрямованому посиленні вже існуючих когнітивних упереджень цільової аудиторії. Такий підхід забезпечує вищу стійкість і тривалість ефекту навіть за умов активної протидії. Акцент робиться на операційно-аналітичному рівні — розглядаються конкретні механізми експлуатації когнітивних, емоційно-афективних та соціально-групових викривлень у реальних інформаційних кампаніях гібридної війни.

– Техніки, що експлуатують когнітивні викривлення — Ця категорія технік спрямована на спотворення процесу обробки, інтерпретації та оцінювання інформації.

– Техніка «Якір + підтверджувальне упередження». Перше повідомлення або теза виконує роль когнітивного якоря, після чого вся наступна інформація інтерпретується крізь призму підтвердження початкового наративу. Класичним прикладом є російська інформаційна кампанія початку 2022 року з тезою «Київ за три дні», коли будь-які подальші повідомлення про труднощі ЗСУ автоматично сприймалися як підтвердження неминучої поразки України.

– Техніка селективного висвітлення. Систематичне поширення лише тих фактів, які узгоджуються з цільовим упередженням, з одночасним ігноруванням або дискредитацією суперечливих даних. Яскравим прикладом є кампанії щодо корупції в Україні, де акцентуються окремі скандали, водночас ігноруються системні антикорупційні реформи та досягнення.

– Техніки, що експлуатують емоційно-афективні викривлення — Дана група технік орієнтована на активацію сильних емоцій, які блокують або суттєво послаблюють раціональний аналіз.

– Техніка емоційного зараження через вірусний контент. Використання коротких, емоційно насичених матеріалів (меми, відеоролики, особисті історії), що максимально активізують афективну

евристику та negativity bias. Прикладом є кампанії з «плачучими матерями» та «м'ясними штурмами», спрямовані на провокування страху, жалю та деморалізації.

– Техніка «Страх + Безнадійність». Створення наративів про неминучість поразки, зраду союзників та марність опору. Ця техніка поєднує афективну евристику з ефектом навченої безпорадності (learned helplessness). Особливо інтенсивно застосовувалася у 2023–2025 роках у кампаніях «Захід зрадив Україну» та «Війна до останнього українця».

– Техніки, що експлуатують соціально-групові викривлення — Ці техніки спрямовані на розкол суспільства, посилення міжгрупової ворожості та радикалізацію позицій.

– Техніка групової поляризації через ехокамери. Створення та підтримка закритих інформаційних середовищ, у яких початкові погляди учасників поступово радикалізуються. Приклади: штучне протиставлення «ватники vs зрадники», «тиловики vs фронтовики», «влада vs народ».

– Техніка соціального доказу + ефекту приєднання до більшості (Ефект бандвагону). Створення ілюзії масової підтримки певної позиції за допомогою бот-мереж, накруток, фейкових свідчень та штучно розкручених трендів. Використовувалася, зокрема, для демонстрації «масового невдоволення мобілізацією».

Класифікація маніпулятивних технік за типом експлуатованих когнітивних викривлень у інформаційних операціях гібридної війни наведена у табл. 2.6

Таблиця 2.6.

Класифікація маніпулятивних технік за типом експлуатованих когнітивних викривлень у інформаційних операціях гібридної війни

Категорія упереджень	**Маніпулятивна техніка**	**Механізм дії**	**Стратегічна мета в гібридній війні**	**Приклад кампанії (2022–2025 рр.)**
Когнітивні	Якір + підтверджувальне упередження Селективне висвітлення	Задавання початкових рамок сприйняття з подальшою інтерпретацією всіх фактів крізь призму початкового наративу; систематичний відбір лише підтверджуючих даних	Формування спотвореної картини реальності та закріплення вигідного агресору наративу	«Київ за три дні» (лютий 2022); кампанії щодо «тотальної корупції» в Україні
Емоційно-афективні	Емоційне зараження через вірусний контент Техніка «Страх + Безнадійність»	Активація сильних негативних емоцій, що блокують раціональний аналіз та активізують афективну евристику та	Деморалізація населення, зниження мотивації до опору та формування психологічної втоми	Кампанії «Безнадійні штурми», «Плачучі матері», «Захід зрадив Україну» (2023–2025)
Соціально-групові	Групова поляризація через ехокамери Соціальний доказ + Ефект Бандвагона	Створення закритих інформаційних середовищ для радикалізації позицій; штучне формування ілюзії масової підтримки	Розкол суспільства, фрагментація національної єдності та послаблення соціальної згуртованості	Протиставлення «ватники vs зрадники», «тил vs фронт», «влада vs народ»; кампанії проти мобілізації

Запропонована класифікація демонструє, що сучасні маніпулятивні техніки в інформаційних кампаніях гібридної війни не є хаотичними, а являють собою продуману систему експлуатації різних типів когнітивних упереджень. Розуміння цієї системи є важливим для подальшого аналізу еволюції та ефективності російських інформаційних кампаній проти України.

Одним із найбільш показових аспектів гібридної війни є системна та високопрофесійна експлуатація когнітивних упереджень у російських інформаційних кампаніях проти України. На відміну від примітивної пропаганди, сучасні операції будуються на точному психологічному розрахунку: виявленні існуючих упереджень різних соціальних груп і їхньому подальшому посиленні.

можна констатувати, що упередження та маніпулятивні техніки становлять основу сучасних інформаційних кампаній гібридної війни. Їхня взаємодія носить системний, адаптивний і високотехнологічний характер. На відміну від класичної пропаганди, сучасні російські інформаційні операції проти України не стільки нав’язують готові наративи, скільки майстерно експлуатують існуючі когнітивні упередження (підтверджувальне, евристику доступності, емоційне зараження, групову поляризацію, упередження авторитету), перетворюючи їх на потужний інструмент досягнення стратегічних цілей.

Ключовими особливостями цієї взаємодії є:

- Висока адаптивність — маніпулятивні техніки постійно еволюціонують залежно від змін у суспільних настроях, технологічного розвитку та ефективності української протидії.
- Комплексність впливу — поєднання когнітивних, емоційних і соціально-групових упереджень у рамках єдиних операцій.
- Технологічне посилення — використання ШІ, deepfake, алгоритмічної персоналізації та бот-мереж суттєво підвищує точність і масштаб впливу.
- Рефлексивний характер — маніпуляції спрямовані не лише на зміну думок, але й на керування процесом прийняття рішень самою жертвою.

Найефективнішими є техніки, які не суперечать внутрішнім упередженням аудиторії, а посилюють їх, створюючи ілюзію самостійного формування висновків. Саме така стратегія дозволяє підтримувати довготривалий інформаційний тиск навіть за умови значної контрпропаганди

2.3. Дослідження ефективності існуючих методів маніпулятивного впливу в кіберпросторі

Оцінка ефективності маніпулятивного впливу в кіберпросторі належить до найбільш складних і водночас стратегічно важливих завдань сучасних досліджень інформаційної війни. Без надійних даних про реальну дієвість тих чи інших технік неможливо розробити ефективні контрзаходи та оптимізувати ресурси протидії.

Значна частина державних програм протидії дезінформації досі базується на інтуїтивних припущеннях, а не на верифікованих даних, що суттєво знижує їхню результативність.

1. Однією з найбільш впливових теоретичних рамок у цій сфері є модель «Науки фейкових новин», запропонована Lazer D. et al. (2018) [24]. Автори розглядають дезінформацію як комплексний соціально-технологічний феномен і пропонують трирівневу модель аналізу її ефективності:

- Індивідуальний рівень — вивчення впливу на когнітивні процеси (когнітивні упередження, евристики, емоційне зараження);
- Мережевий рівень — аналіз механізмів поширення та вірусності контенту в соціальних мережах;
- Інституційний рівень — оцінка впливу дезінформації на суспільні інститути, політичні процеси та суспільну згуртованість.

Перевага цієї моделі полягає в тому, що вона підкреслює: ефективність маніпулятивного впливу визначається не лише якістю самого контенту, а й взаємодією психологічних особливостей аудиторії, архітектури цифрових платформ та ширшого соціально-політичного контексту.

2. Теорія психологічної імунізації (Inoculation Theory)

Теорія психологічної імунізації, започаткована В. МакГайром у 1960-х роках і суттєво розвинена сучасними дослідниками (Roozenbeek, van der Linden та ін., 2019–2022), є однією з найбільш перспективних наукових моделей протидії дезінформації [25]. Згідно з цією теорією, попереднє ознайомлення

людини з ослабленою формою маніпулятивної техніки («імунізація») формує когнітивну стійкість, аналогічно до дії вакцини, і підвищує опірність до повноцінної дезінформації в майбутньому.

3. Модель психологічних драйверів дезінформації (Ecker et al., 2022)

Ecker U.K.H. et al. (2022) у своєму огляді «The Psychological Drivers of Misinformation Belief and Its Resistance to Correction» запропонували інтегровану модель, яка пояснює механізми стійкості хибних переконань навіть після їх спростування [26]. Автори виділяють чотири основні групи психологічних драйверів:

- Когнітивні — упередження, евристики та обмеженість аналітичного мислення;
- Соціальні — групова ідентичність, соціальний доказ та конформізм;
- Емоційні — афективні реакції та емоційне зараження;
- Мотиваційні — бажання зберегти когнітивну узгодженість світогляду та уникнути когнітивного дисонансу.

Ця модель має велике теоретичне та практичне значення, оскільки дозволяє не лише оцінювати силу маніпулятивного впливу, але й прогнозувати стійкість його наслідків до корекції.

Для об’єктивної оцінки ефективності методів маніпуляції доцільно використовувати багаторівневу систему критеріїв, яка враховує як часовий вимір (коротко-, середньо- та довгостроковий ефект), так і характер наслідків (когнітивні, поведінкові та соціально-політичні). Такий підхід дозволяє комплексно оцінити як безпосередній вплив маніпуляцій, так і їхню стійкість та ширші суспільні наслідки (табл. 2.7)

Таблиця 2.7.

Критерії оцінки ефективності маніпулятивного впливу

Рівень ефективності	**Критерій оцінки**	**Показники вимірювання**
Короткостроковий	Охоплення та швидкість поширення	Кількість переглядів, репостів, охоплення аудиторії, швидкість вірусності
Середньостроковий	Зміна переконань та установок	Динаміка відповідей у соціологічних опитуваннях, зміна рівня згоди з наративом
Довгостроковий	Стійкість переконань	Збереження віри в дезінформацію після спростування (continued influence effect)
Поведінковий	Зміна реальної поведінки	Участь у протестах, підтримка певних політичних рішень, бойкоти, ухилення від мобілізації
Соціально-політичний	Суспільні та політичні наслідки	Рівень суспільної поляризації, динаміка довіри до державних інститутів, соціальна згуртованість

Комплексний підхід до оцінки ефективності маніпулятивного впливу є необхідним, оскільки кожна з розглянутих теоретичних моделей має певні обмеження.

У зв’язку з цим обрано інтегрований комплексний підхід, який поєднує елементи всіх трьох моделей. Такий синтез дозволяє провести повноцінний багаторівневий аналіз — від механізмів поширення та когнітивної обробки інформації до довгострокових соціально-політичних наслідків, — з урахуванням специфіки українського інформаційного простору в умовах триваючої гібридної війни.

Фактчекінг та реактивне спростування дезінформації залишаються одними з найбільш поширених контрзаходів у боротьбі з маніпулятивним впливом. Однак їхня реальна ефективність є предметом гострих наукових дискусій. Сучасні дослідження свідчать, що реактивні методи протидії не є універсально ефективними і в певних умовах можуть навіть посилювати ефект маніпуляції.

Chan M. et al. (2023) у своєму мета-аналізі, який охопив понад 50 експериментальних досліджень, встановили, що фактчекінг у середньому знижує віру в дезінформацію на 0,19–0,25 стандартного відхилення. Цей ефект є статистично значущим, але відносно modest. Більш важливим є висновок авторів про те, що ефективність фактчекінгу суттєво варіюється залежно від контексту та характеристик аудиторії [27].

Walter N. та Murphy S.T. (2018) у своєму мета-аналізі виявили, що хоча спростування зменшує вплив дезінформації, воно рідко повністю усуває його. У середньому після спростування зберігається близько 40–60% первинного ефекту фейкової інформації [28].

Ефективність фактчекінгу суттєво залежить від таких умов:

1. Своєчасність — спростування, надане протягом перших 24–48 годин, працює значно краще.
2. Джерело — спростування від нейтральних або довірених джерел є ефективнішим, ніж від влади.
3. Формат — поєднання тексту з візуальними елементами підвищує ефективність на 15–25%.
4. Повторюваність — одноразове спростування має слабкий ефект; необхідне системне повторення.

Превентивні стратегії, засновані на теорії імунізації демонструють значно кращі результати у довгостроковій перспективі, на відміну від реактивних методів, превентивні працюють проактивно: вони не спростовують конкретну брехню, а навчають розпізнавати маніпулятивні техніки. Це робить їх більш універсальними та стійкими до нових наративів.

Фактчекінг та реактивне спростування дезінформації мають обмежену ефективність і повинні розглядатися як допоміжний, а не основний інструмент протидії. Найперспективнішим напрямком є поєднання превентивних стратегій (психологічна імунізація) з селективним використанням якісного фактчекінгу. Для України, яка перебуває під постійним інформаційним тиском, пріоритетом

має стати розвиток превентивних програм, оскільки реактивні методи не здатні повністю компенсувати шкоду від масованого маніпулятивного впливу.

Превентивні стратегії протидії маніпулятивному впливу, засновані на теорії психологічної імунізації), на сьогодні вважаються одним із найбільш перспективних напрямків у сфері когнітивної безпеки. На відміну від реактивних методів (фактчекінг і спростування), превентивні стратегії спрямовані не на корекцію вже сформованих хибних переконань, а на підвищення стійкості когнітивної системи до майбутніх маніпуляцій.

Melisa Basol в своїй статті «Good News about Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News »підтвердила, що гейміфіковані методи імунізації є особливо ефективними для молодої аудиторії, яка становить основну цільову групу для сучасних маніпуляцій у кіберпросторі [29].

Гейміфікація демонструє значні переваги порівняно з традиційними освітніми програмами. Ігрові формати підвищують залученість учасників на 60–80% порівняно зі звичайними лекціями чи статтями. Крім того, гейміфікація дозволяє моделювати реальні механізми поширення дезінформації, що сприяє кращому засвоєнню матеріалу.

Превентивні стратегії демонструють суттєву перевагу над реактивними в кількох ключових аспектах:

- Стійкість до повторних атак — імунізовані учасники краще протистоять новим маніпуляціям, навіть якщо вони відрізняються від тих, на яких їх імунізували.
- Масштабованість — освітні програми та ігри можна впроваджувати системно в школах та університетах.
- Відсутність зворотнього ефекту — на відміну від спростування, превентивні методи майже не викликають захисної реакції.

Водночас превентивні стратегії мають певні обмеження: вони вимагають часу та системної роботи, їхній ефект проявляється не відразу, а через певний період, і вони менш ефективні проти вже закріплених глибоких переконань.

Для України, яка перебуває під постійним і масованим інформаційним тиском, розвиток превентивних стратегій є стратегічно пріоритетним. Комбінація системної медіаграмотності з елементами психологічної імунізації може стати одним із найбільш ефективних інструментів підвищення когнітивної стійкості суспільства в умовах тривалої гібридної війни.

Превентивні стратегії, особливо гейміфіковані форми психологічної імунізації, демонструють значно вищу довгострокову ефективність порівняно з реактивними методами. Їхнє системне впровадження може стати ключовим елементом національної стратегії когнітивної безпеки України.

Ефективність маніпулятивних методів у кіберпросторі не є константою і значною мірою залежить від комплексу контекстуальних, технологічних, соціально-психологічних та культурних факторів. Розуміння цих факторів дозволяє перейти від загальних оцінок до більш точного прогнозування успішності інформаційних операцій у конкретних умовах, зокрема в українському інформаційному просторі.

1. Контекстуальні фактори

Найсуттєвішим контекстуальним фактором є кризовий період. Zhou Y. (2024) у своєму дослідженні доводить, що під час гострих криз (військові дії, енергетичні проблеми, політична нестабільність) ефективність маніпуляцій, заснованих на емоційному зараження та евристиці доступності, зростає на 40–65%. У стабільних умовах цей показник суттєво знижується [30].

Рівень суспільної тривожності також є потужним модифікатором. У періоди високої тривожності (2022 рік) навіть примітивні наративи демонстрували високу ефективність. У 2024–2025 роках, зі зниженням рівня гострої тривоги, ефективність змістилася в бік більш витончених, довгострокових технік (підтверджувальне упередження та рефлексивне управління).

2. Технологічні фактори

Алгоритмічна архітектура платформ є одним із ключових посилювачів маніпуляцій. Рекомендаційні алгоритми соціальних мереж створюють умови для посилення підтверджувального упередження та групової поляризації [31].

Рівень довіри до інститутів є одним із найбільш впливових факторів. У суспільствах з низькою довірою до влади (Україна 2022–2025) ефективність дискредитуючих кампаній значно вища.

Групова ідентичність також відіграє критичну роль. Чим сильнішою є ідентифікація людини з певною групою, тим важче спростувати інформацію, яка захищає цю ідентичність.

3. Культурні та національні особливості українського інформаційного простору

Український інформаційний простір має низку специфічних характеристик, які впливають на ефективність маніпуляцій:

- Високий рівень горизонтальної довіри (довіра між людьми) при відносно низькій вертикальній (довіра до інститутів).
- Значна регіональна та мовна диференціація, яка створює природні умови для сегментованої маніпуляції.
- Висока адаптивність населення до інформаційного тиску (ефект «пристосування» після 2014 року).
- Сильна емоційна реактивність під час воєнних подій.

Для українського інформаційного простору характерна підвищена вразливість до емоційно-заряджених і дискредитуючих технік у кризові періоди, водночас спостерігається зростання стійкості завдяки адаптивності населення.

Ефективність маніпулятивних методів у кіберпросторі є результатом складної взаємодії контекстуальних, технологічних, соціально-психологічних та культурних факторів. Для українського інформаційного простору характерна підвищена вразливість до емоційно-заряджених і дискредитуючих технік у кризові періоди. Детальне відображення основних факторів, які посилюють ефективність маніпулятивного впливу в українському інформаційному просторі

наведені на рис. 2.1.

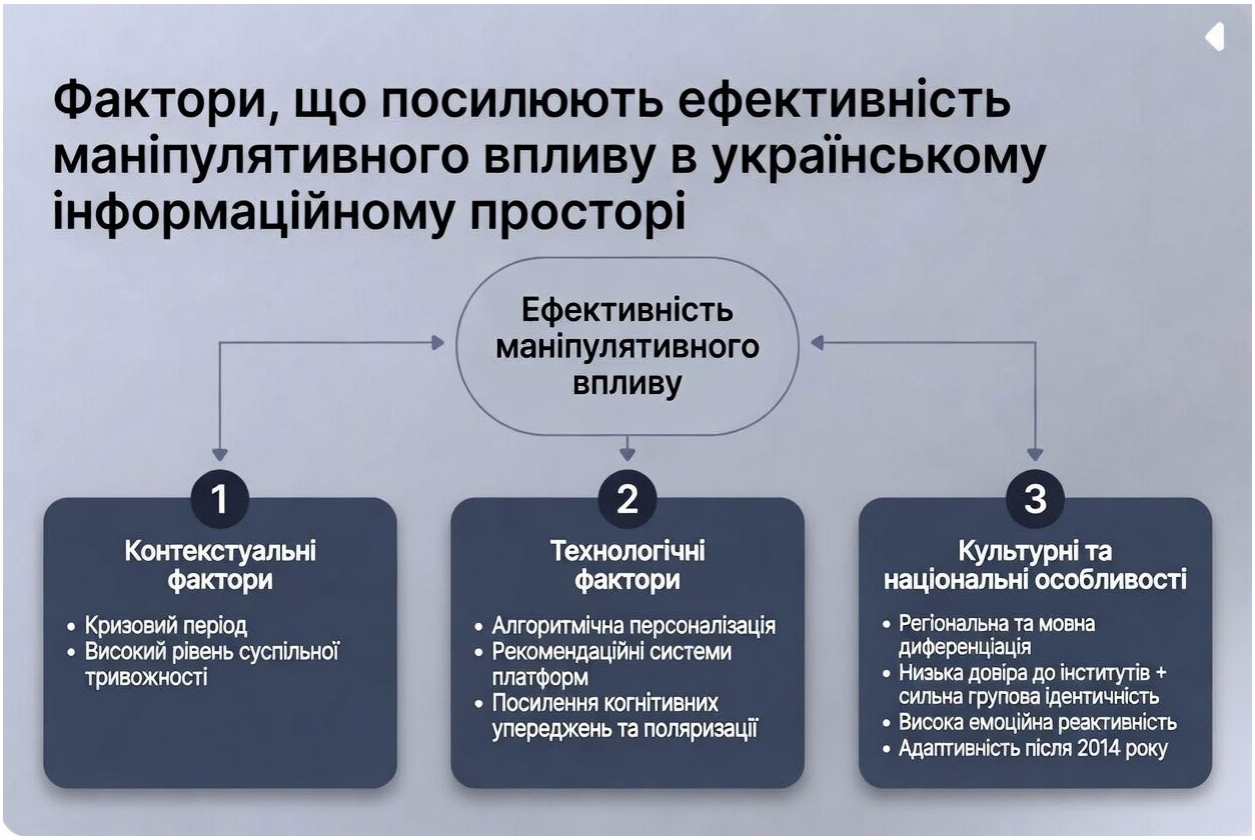

Рис. 2.1 Фактори, що посилюють ефективність маніпулятивного впливу в українському інформаційному просторі

Методи, засновані на експлуатації когнітивних упереджень, демонструють різну ефективність залежно від контексту. Найвищу результативність показують комбіновані техніки, які поєднують підтверджувальне упередження з емоційним зараженням та груповою поляризацією. Ці методи особливо ефективні в кризові періоди та в умовах алгоритмічної персоналізації, що характерно для російських ІПСО проти України

Реактивні методи протидії (фактчекінг і спростування) мають обмежену ефективність. Вони здатні частково зменшити вплив дезінформації в короткостроковій перспективі, але демонструють слабку стійкість і ризики backfire effect. У довгостроковій перспективі вони значно поступаються превентивним стратегія

Превентивні стратегії, зокрема психологічна імунізація (inoculation) та гейміфіковані програми медіаграмотності, є найбільш перспективними. Вони

забезпечують вищу стійкість до маніпуляцій, кращу довгострокову ефективність і нижчий ризик негативних ефектів

Ефективність маніпулятивного впливу визначається не лише якістю самого контенту, а й складною взаємодією психологічних, технологічних та контекстуальних факторів. У випадку України це створює ситуацію, коли традиційні методи протидії вже недостатні для забезпечення когнітивної безпеки суспільства.

Висновки до розділу 2

У цьому розділі проведено аналіз когнітивних викривлень як інструменту маніпулятивного впливу в сучасних інформаційно-психологічних операціях, зокрема в умовах гібридної війни.

Було встановлено, що когнітивні викривлення є одним із ключових елементів реалізації ІПСО. Розглянуто механізми їх використання Російською Федерацією проти України на різних етапах війни. Показано, що ефективність таких операцій ґрунтується на системній сегментації цільових аудиторій, адаптації наративів під домінуючі когнітивні уразливості та поєднанні різних типів викривлень (підтверджувальне упередження, емоційне зараження, евристика доступності тощо).

Також здійснено систематизацію маніпулятивних технік, що експлуатують когнітивні викривлення. Виділено три основні категорії технік (когнітивні, емоційно-афективні та соціально-групові) та проаналізовано їх взаємодію з теоретичними моделями впливу.

В кінці проведено оцінку ефективності методів маніпулятивного впливу та існуючих підходів до протидії їм. Виявлено обмежену ефективність реактивних методів (фактчекінг, спростування) через ефект продовженого впливу та зовортій ефект дії. Водночас підтверджено високу перспективність превентивних стратегій, зокрема теорії психологічної імунізації та

гейміфікованих програм. Проаналізовано фактори, що впливають на ефективність маніпуляцій в українському інформаційному просторі.

Розділ 3 РОЗРОБКА МЕТОДІВ ТА РЕКОМЕНДАЦІЙ ЩОДО ЗАХИСТУ ВІД КОГНІТИВНИХ ЗАГРОЗ

3.1. Технічні та програмні засоби детекції маніпулятивного контенту

У сучасних умовах гібридної війни, де інформаційно-психологічні операції (ІПСО) стали одним із основних інструментів впливу, розробка ефективних технічних засобів детекції маніпулятивного контенту набуває стратегічного значення.

Когнітивно орієнтовані маніпуляції, спрямовані на експлуатацію психологічних уразливостей аудиторії, поширюються з безпрецедентною швидкістю та масштабом, що робить традиційні методи моніторингу недостатніми.

Актуальність проблеми зумовлена радикальною зміною характеру інформаційного протиборства. Якщо раніше основну загрозу становила пряма пропаганда, то сьогодні переважна більшість маніпуляцій будується на тонкій експлуатації когнітивних упереджень, емоційного зараження та алгоритмічної персоналізації [31].

В умовах повномасштабної російської агресії інформаційний простір України зазнає постійного високотехнологічного тиску. Існуючі інструменти детекції демонструють суттєві обмеження. Більшість комерційних і відкритих рішень (наприклад, Google Fact Check Tools, Meta’s AI Classifier) орієнтовані переважно на виявлення класичної дезінформації (фактично неправдиві твердження), але значно гірше справляються з виявленням когнітивно орієнтованого маніпулятивного контенту, який може бути формально правдивим, але спрямованим на активацію упереджень, емоційне зараження чи групову поляризацію.

Сучасні системи детекції мають низьку ефективність при аналізі мультимодального контенту (відео + текст + емоційний контекст) та персоналізованих наративів. Крім того, більшість інструментів розроблено для англомовного простору і показує знижену точність при роботі з

російськомовним та українськомовним контентом через особливості мови, культурного контексту та стилістики маніпуляцій.

Для України проблема посилюється кількома специфічними факторами:

- домінуванням Telegram як основної платформи поширення маніпуляцій;
- високою швидкістю поширення контенту в умовах війни;
- поєднанням державного, квазідержавного та приватного маніпулятивного контенту;
- необхідністю одночасного аналізу як російськомовного, так і українськомовного сегментів.

Розробка адаптованих технічних та програмних засобів детекції маніпулятивного контенту є не лише технічним, а й стратегічним завданням національної безпеки. Існуючі інструменти не в повній мірі відповідають специфіці когнітивних загроз, з якими стикається Україна [32]. Це зумовлює необхідність створення гібридних систем, здатних виявляти не лише фактологічну неправду, але й психологічно орієнтовані маніпуляції. Аналіз теоретичних основ, класифікації та перспективних технологій дозволить сформувати концептуальні засади для розробки таких рішень.

Сучасні методи автоматичної детекції маніпулятивного контенту можна класифікувати за чотирма ключовими парадигмами. У таблиці нижче наведено їх порівняльну характеристику.

Під рівнем зрілості розуміється ступінь розробленості, емпіричної апробації та практичного застосування відповідного підходу в реальних системах детекції:

Високий рівень зрілості вказує на наявність численних перевірених рішень, відпрацьованих алгоритмів і широкого використання в промислових та наукових системах.

Середній–високий рівень свідчить про активний розвиток технології з перспективними результатами, але ще обмеженою кількістю масштабних реалізацій.

Таблиця 3.1.

Порівняльна характеристика основних парадигм детекції маніпулятивного контенту

Парадигма	**Основний об'єкт аналізу**	**Сильні сторони**	**Обмеження**	**Рівень зрілості**
Лінгвістичний підхід	Стилістичні, лексичні та синтаксичні особливості тексту	Висока інтерпретованість результатів, добре виявляє класичні пропагандистські маркери (емоційна лексика, повторення, риторичні прийоми)	Слабко працює з мультимодальним контентом, сарказмом та імпліцитними маніпуляціями	Високий
Семантичний підхід	Смисловий зміст тексту (NLP, BERT, RoBERTa, векторні моделі)	Виявлення прихованих смислових маніпуляцій,	Висока ресурсоємність, потребує великих обсягів якісних даних для навчання	Високий
Поведінковий підхід	Патерни поширення, взаємодія користувачів, аномалії поведінки акаунтів	Ефективне виявлення бот-мереж, coordinated inauthentic behavior та штучного посилення наративів	Складність інтерпретації причинності, залежність від якості даних про взаємодію	Середній–високий
Мережевий (графовий) підхід	Поширення інформації як графова структура (центральність вузлів, кластери, шляхи дифузії)	Добре виявляє організовані кампанії, джерела та шляхи поширення маніпуляцій	Висока обчислювальна складність, проблеми з аналізом в реальному часі	Середній–високий

Сучасні системи детекції маніпулятивного контенту розвиваються в рамках кількох ключових концепцій, які відрізняються об'єктом аналізу, методологічним апаратом та рівнем практичної зрілості [33]. Під рівнем зрілості у таблиці розуміється ступінь розробленості та практичної готовності відповідної концепції до широкого застосування. Високий рівень свідчить про наявність численних перевірених рішень і промислових реалізацій. Середній

рівень вказує на активний розвиток технології з перспективними результатами, але обмеженою кількістю масштабних впроваджень. Низький–середній рівень характерний для новітніх напрямків, які мають високий потенціал, але потребують подальших досліджень і технічної доопрацювання.

Таблиця 3.2.

Порівняння ключових концепцій детекції маніпулятивного контенту

Концепція	**Основний об'єкт виявлення**	**Сильні сторони**	**Обмеження**	**Рівень зрілості**
Детекція фейкових новин	Фактологічна неправда	Висока точність на простих фейках, добре відпрацьовані алгоритми	Погано працює з «правдивими маніпуляціями»	Високий
Детекція пропаганди	Техніки пропаганди	Добре виявляє стилістичні та риторичні маркери пропаганди	Слабка адаптивність до культурного контексту та імпліцитних маніпуляцій	Середній
Когнітивна детекція маніпуляцій	Психологічний вплив (активація когнітивних упереджень)	Найвища релевантність для сучасних ІПСО, фокус на прихованому маніпулятивному намірі	Найскладніший у реалізації, потребує інтеграції психологічних моделей	Низький–середній

Під рівнем зрілості у таблиці розуміється ступінь розробленості та практичної готовності відповідної концепції до широкого застосування.

– Високий рівень свідчить про наявність численних перевірених рішень і промислових реалізацій.

- Середній рівень вказує на активний розвиток технології з перспективними результатами, але обмеженою кількістю масштабних впроваджень.
- Низький–середній рівень характерний для новітніх напрямків, які мають високий потенціал, але потребують подальших досліджень і технічної доопрацювання.

Сучасні засоби детекції маніпулятивного контенту демонструють значну різноманітність підходів. Найперспективнішим напрямком є розвиток гібридних систем, які поєднують мультимодальний контентний аналіз з мережевим і контекстним рівнями. Для України особливо важливим є створення національних рішень, адаптованих до специфіки російськомовного маніпулятивного контенту, домінування Telegram та когнітивно орієнтованих технік впливу.

Існуючі інструменти детекції маніпулятивного контенту демонструють помірну ефективність при виявленні класичної дезінформації, але суттєво поступаються у виявленні складних когнітивно орієнтованих маніпуляцій. Це зумовлює нагальну потребу в розробці спеціалізованих гібридних систем, адаптованих до специфіки гібридної війни та українського інформаційного середовища [34]. Подальший розвиток повинен бути спрямований на поєднання потужних мовних моделей з психологічними індикаторами маніпуляції та мережевим аналізом.

Розробка національної системи детекції маніпулятивного контенту є стратегічно необхідним кроком для забезпечення когнітивної безпеки населення України в умовах триваючої гібридної війни. Існуючі фрагментарні рішення не здатні ефективно протистояти масштабним, адаптивним і високотехнологічним інформаційно-психологічним операціям.

Саме тому пропонується створення інтегрованої національної системи когнітивної детекції (National Cognitive Threat Detection System — NCTDS) — комплексного гібридного рішення, яке поєднує методи штучного інтелекту, психологічні моделі та мережевий аналіз для виявлення не лише фактологічної

дезінформації, а й контенту, що експлуатує когнітивні викривлення. Архітектура запропонованої системи представлена на рис. 3.1.

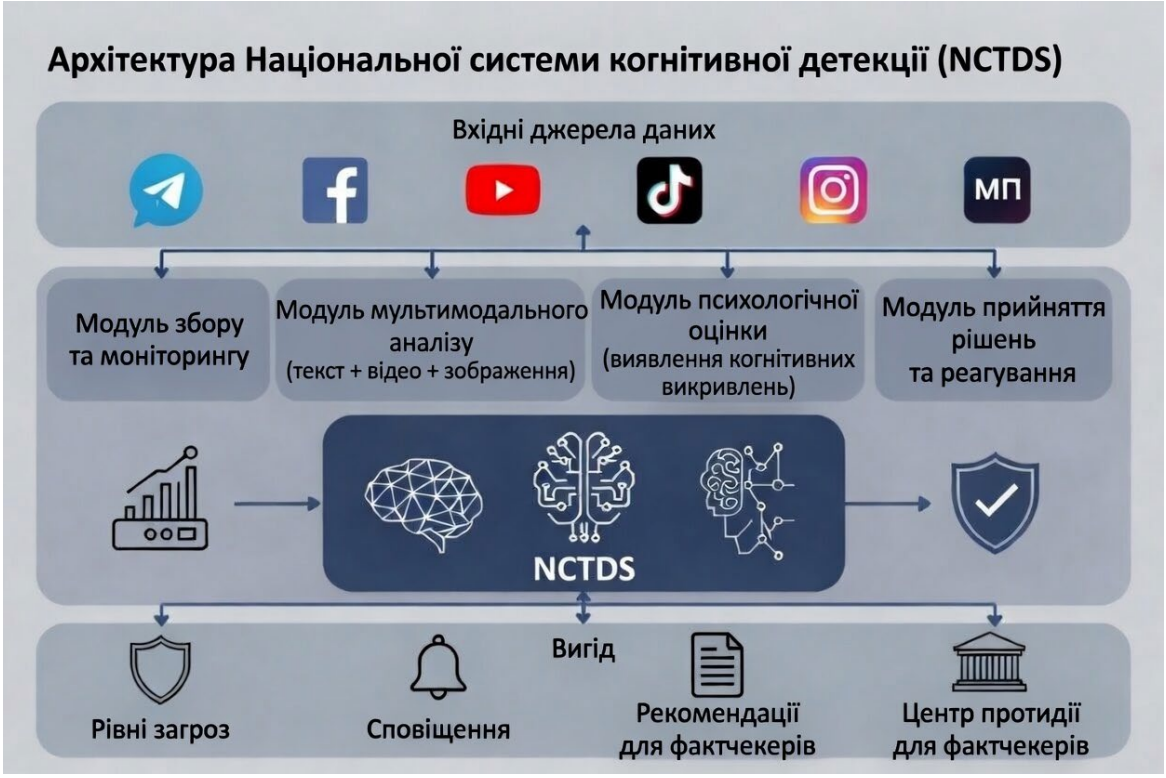

Рис. 3.1 Архітектура Національної системи когнітивної детекції (NCTDS)

Основними принципами функціонування Національної системи когнітивної детекції (NCTDS) є:

- Мультимодальність. Система здійснює одночасний аналіз різних типів контенту — текстового, візуального, аудіовізуального та метаданих. Це дозволяє виявляти маніпулятивні техніки, які можуть бути приховані в одному форматі, але чітко проявляються при комплексному розгляді (наприклад, емоційно заряджений deepfake з нейтральним текстовим описом).
- Гібридність. NCTDS поєднує переваги штучного інтелекту (великі мовні моделі, комп'ютерний зір, графовий аналіз) з експертним знанням людини. Автоматизовані моделі виконують первинну обробку великих обсягів даних, тоді як психологи, фактчекери та аналітики здійснюють фінальну

верифікацію критичних випадків. Такий підхід суттєво підвищує точність і надійність виявлення.

- Адаптивність. Система постійно навчається на нових даних українського інформаційного простору. Завдяки механізмам онлайн-навчання вона швидко адаптується до зміни тактик маніпуляцій, появи нових наративів та еволюції мови ворога. Це особливо важливо в умовах динамічної гібридної війни.
- Проактивність. На відміну від традиційних систем фактчекінгу, які реагують уже на поширену дезінформацію, NCTDS орієнтована на раннє виявлення потенційно небезпечних інформаційних кампаній. Система аналізує патерни поширення, швидкість зростання охоплення та психологічний потенціал контенту ще до того, як він набуде масового характеру.

Система складається з чотирьох основних взаємопов'язаних модулів, які працюють у єдиному конвеєрі обробки даних, а загальний процес обробки інформації в NCTDS відбувається наступним чином: дані послідовно проходять через усі модулі, при цьому на кожному етапі відбувається збагачення інформації та підвищення точності аналізу. Такий підхід дозволяє ефективно поєднувати швидкість автоматизованої обробки з глибиною психологічного та контекстного аналізу.

Основними складовими архітектури системи є:

1. Модуль збору та моніторингу — здійснює постійний збір даних з основних джерел українського інформаційного простору (Telegram-канали, соціальні мережі, новинні сайти, YouTube та ін.). Модуль працює в режимі реального часу та забезпечує первинну фільтрацію контенту за заданими критеріями.

Для реалізації модуля обрано гібридний підхід, що поєднує спеціалізовані інструменти для різних типів платформ:

1. Для Telegram-каналів та чатів використовуються бібліотеки Telethon та Pyrogram. Ці інструменти дозволяють здійснювати повноцінне API-

взаємодію з Telegram (MTProto-протокол), отримувати повідомлення в реальному часі через функцію long polling, збирати метадані (кількість переглядів, репостів, реакцій), а також працювати з приватними каналами та супергрупами. Telethon забезпечує високу стабільність і низький рівень блокувань, тоді як Pyrogram пропонує зручніший асинхронний інтерфейс для високонавантажених систем.

2. Для інших платформ (Facebook, Instagram, YouTube, Twitter/X, новинні сайти) застосовується комбінація Scrapy (фреймворк для веб-скрапінгу) та Selenium WebDriver. Scrapy забезпечує високу швидкість і ефективність при зборі структурованих даних з відкритих джерел, а Selenium використовується для динамічного контенту, що завантажується через JavaScript (сторіси, рілси, live-трансляції), а також для обходу простих систем анти-бот захисту.
3. Для забезпечення масштабованості, відмовостійкості та оркестрації модуль розгортається у контейнеризованому середовищі з використанням Docker та Kubernetes. Кожен тип збирача (Telegram-collector, Web-scraper тощо) працює в окремих подах, що дозволяє горизонтальне масштабування залежно від навантаження. Оркестрація через Kubernetes забезпечує автоматичне відновлення після збоїв, балансування навантаження та моніторинг ресурсів (Prometheus + Grafana).

Додатковими важливими елементами модуля є:

– Система черг завдань (Celery + Redis/RabbitMQ) для асинхронної обробки великих обсягів даних;

– Розумна фільтрація на етапі збору (pre-filtering) для зменшення навантаження на наступні модулі;

– Зберігання сирих даних у розподіленому сховищі (Apache Kafka або MinIO) для подальшого ретроспективного аналізу.

2. Модуль мультимодального аналізу — є одним із центральних компонентів Національної системи когнітивної детекції (NCTDS), оскільки сучасний маніпулятивний контент у кіберпросторі переважно має

мультимодальний характер. Він поєднує текст, зображення, відео, аудіо та метадані, що значно ускладнює виявлення прихованих когнітивних маніпуляцій. Метою модуля є комплексна семантико-психологічна обробка контенту різних модальностей з метою виявлення ознак маніпулятивного впливу, спрямованого на експлуатацію когнітивних викривлень. Модуль реалізує багаторівневий аналіз, що включає незалежну обробку кожної модальності з подальшою їхньою інтеграцією) для формування єдиного векторного представлення контенту.

Для забезпечення високої точності та адаптації до українсько-російського мовного середовища в модулі застосовується спеціалізований технологічний стек:

1. Використовується багатофункціональна мовна модель на базі multilingual BERT або ukr-roberta, доповнена fine-tuned версією XLM-RoBERTa. Модель попередньо донавчалася на спеціально розміченому корпусі українського та російського маніпулятивного контенту (включаючи наративи ІПСО 2022–2025 рр.).

Основні завдання текстового аналізу:

- Виявлення стилістичних маркерів пропаганди (емоційна лексика, риторичні прийоми, повторення).
- Оцінка наявності когнітивних «якорів» та поляризуючої лексики.
- Розпізнавання імпліцитних маніпуляцій (сарказм, евфемізми, рамкування).

2. Для візуальної складової застосовуються моделі CLIP (OpenAI) або SigLIP, які дозволяють ефективно поєднувати візуальне та текстове представлення. Моделі fine-tuned на задачі виявлення емоційного забарвлення та маніпулятивних візуальних прийомів (шокуючі образи, селективне кадрування, меміфікація тощо).

Ключові метрики:

- Емоційне навантаження зображення.
- Відповідність між візуальним та текстовим повідомленням (виявлення невідповідності як маркера маніпуляції).

– Виявлення ознак візуальної пропаганди (стереотипізація, героїзація/демонізація).

3. Для відеоконтенту використовуються трансформерні архітектури VideoMAE та TimeSformer, які демонструють високу ефективність у задачах виявлення діпфейків. Додатково інтегрується бібліотека DeepFace (або її розширені варіанти) для аналізу емоцій обличчя, мікроекспресій та синхронізації аудіо-візуальних даних.

Основні функції відеомодуля:

– Детекція діпфейків та технологічно згенерованого контенту.

– Аналіз динаміки емоційного впливу (зміна емоційного забарвлення протягом відео).

– Виявлення технік емоційного зараження через візуальні стимули.

4. Після незалежної обробки кожної модальності відбувається late fusion — об’єднання отриманих векторних представлень через багатошаровий персептрон або трансформерний ф’южн-модуль. Це дозволяє моделі враховувати взаємодію між модальностями (наприклад, нейтральний текст + шокуюче зображення = високий маніпулятивний потенціал).

5. Для підвищення інтерпретовності результатів застосовується SHAP (SHapley Additive exPlanations) або LIME, що дозволяє пояснювати, яка саме модальність та які конкретні ознаки найбільше вплинули на оцінку загрози.

3. Модуль психологічної оцінки — ключовий модуль системи, який здійснює оцінку контенту з точки зору його потенціалу щодо експлуатації когнітивних викривлень. Основною метою модуля є кількісна та якісна оцінка ступеня психологічного впливу контенту на цільову аудиторію через активацію систематичних когнітивних спотворень. Модуль аналізує наявність і силу прояву таких механізмів, як підтверджувальне упередження, емоційне зараження, групова поляризація, ефект ехокамери, евристика доступності, ефект якоря, негативне упередження та інші [35].

Для реалізації модуля психологічної оцінки застосовується гібридний підхід, що поєднує методи машинного навчання з елементами промпт-інжинірингу та психолінгвістичного аналізу:

1. Кастомна класифікаційна модель

– Пропонується використовувати спеціалізовану модель ukr-roberta-base на базі fine-tuned великих мовних моделей, навчених на унікальному датасеті українського та російського маніпулятивного контенту. Датасет містить тисячі вручну розмічених прикладів з експертним маркуванням за типами когнітивних викривлень. Модель здатна одночасно виконувати багатозадачну класифікацію та регресію, оцінюючи ймовірність та інтенсивність активації кожного типу викривлення.

2. Промпт-інжиніринг та LLM-асистент

– Доповненням до нейронної моделі слугує система на базі відкритих великих мовних моделей Llama-3 або Gemma-2. За допомогою промптів модель проводить семантичний і психологічний розбір контенту.

– Промпти включають чіткі психологічні критерії та приклади маркування, що суттєво підвищує інтерпретованість результатів.

3. Психолінгвістичні та стилометричні індикатори

– Модуль автоматично розраховує комплекс психологічних ознак, зокрема:

- Кількість та інтенсивність емоційно-заряджених слів і виразів;
- Наявність і силу когнітивних «якорів»;
- Рівень поляризуючої та дихотомічної лексики («свої» та «чужі», «зрада» та «героїзм»);
- Ступінь використання технік соціального доказу та ефекту більшості;
- Індекс емоційного зараження (на основі лексичного та синтаксичного аналізу);
- Маркери групової ідентифікації та інших соціально-групових викривлень.

На виході модуль формує психологічний профіль контенту — вектор оцінок за ключовими когнітивними викривленнями, а також загальний індекс маніпулятивного потенціалу (Cognitive Manipulation Score, CMS) у діапазоні 0–100.

Запровадження модуля психологічної оцінки дозволяє перейти від фактологічного до когнітивно-орієнтованого підходу в детекції інформаційних загроз. Такий модуль здатен виявляти контент, який формально не містить відвертої брехні, проте є високоефективним інструментом маніпуляції свідомістю. Це особливо важливо в умовах сучасної гібридної війни, де противник активно використовує «правдиві маніпуляції»

4. Модуль прийняття рішень та реагування — є завершальним компонентом конвеєра обробки Національної системи когнітивної детекції (NCTDS). Він виконує функцію інтеграції та інтерпретації результатів усіх попередніх модулів, перетворюючи багатовимірні дані аналізу на конкретні оцінки загрози та операційні рекомендації. Саме цей модуль забезпечує перехід від чисто аналітичної фази до практичних дій, роблячи систему не лише детектувальною, а й дієвою інструментом когнітивної безпеки.

Основними завданнями модуля є:

– Комплексна оцінка рівня загрози контенту;

– Генерація рекомендацій щодо оптимальних заходів реагування;

– Автоматизоване виконання низькоприоритетних дій;

– Забезпечення механізму зворотного зв’язку для безперервного вдосконалення системи.

Для реалізації модуля обрано гібридний підхід, що поєднує методи машинного навчання з логікою на правила (rule-based). Основним інструментом є градієнтний бустинг над деревами рішень — XGBoost та CatBoost. Ці алгоритми демонструють високу ефективність у задачах ранжування та класифікації при роботі з гетерогенними даними, що надходять з мультимодального та психологічного модулів.

Вхідними даними для модуля є:

– Векторні представлення контенту з модуля мультимодального аналізу;

– Психологічний профіль контенту та індекс Cognitive Manipulation Score (CMS) з модуля психологічної оцінки;

– Додаткові контекстні ознаки (джерело, швидкість поширення, охоплення аудиторії, сегмент цільової групи тощо).

На основі цих даних модуль обчислює інтегральний рівень загрози за шкалою від 0 до 100, який поділяється на чотири рівні:

– Низький (0–35) — моніторинг;

– Середній (36–65) — автоматичне маркування та пріоритетне сповіщення;

– Високий (66–85) — негайне сповіщення фактчекерів та модераторів;

– Критичний (86–100) — автоматична передача матеріалів до державних органів (Центр протидії дезінформації, СБУ, Мінцифри тощо).

– Модуль використовує комбінацію предиктивної моделі та набору експертних правил для формування рекомендацій. Для кожного виявленого загрозливого контенту система генерує персоналізований набір дій, зокрема:

– Автоматичне маркування контенту тегами (тип викривлення, цільова аудиторія, ступінь небезпеки);

– Сповіщення фактчекінгових організацій через API;

– Підготовка та передача матеріалів до уповноважених державних органів;

– Рекомендації щодо пре-банкінгу або контрнаративів для публічного реагування;

– Блокування або обмеження поширення в разі інтеграції з платформами (за наявності відповідних домовленостей).

Важливою особливістю модуля є вбудований механізм зворотного зв’язку. Після виконання рекомендацій та отримання експертної верифікації (від фактчекерів або державних фахівців) результати повертаються в систему як нові розмічені дані. Це дозволяє здійснювати регулярне донавчання всіх попередніх

модулів, підвищуючи точність і адаптивність NCTDS до еволюції тактик противника.

Для управління процесом навчання застосовується активне навчання — система автоматично виділяє найбільш неоднозначні випадки для пріоритетної експертної розмітки, що оптимізує використання обмежених людських ресурсів.

Запровадження модуля прийняття рішень та реагування дозволяє реалізувати повний цикл когнітивної протидії: від виявлення загрози до оперативного реагування та вдосконалення системи [36]. Гібридна архітектура забезпечує баланс між швидкістю автоматизованих рішень і надійністю експертного контролю, мінімізуючи ризик помилкових позитивних результатів.

Така чотирьохмодульна архітектура дозволяє забезпечити повний цикл обробки інформації — від збору сирих даних до вироблення конкретних рекомендацій щодо реагування, що є критично важливим в умовах швидкоплинних інформаційно-психологічних операцій.

Запропонована архітектура Національної системи когнітивної детекції (NCTDS) має ряд суттєвих переваг порівняно з існуючими інструментами протидії дезінформації. Насамперед вона дозволяє здійснити перехід від реактивного підходу (фактчекінг уже поширеного контенту) до проактивного виявлення когнітивних загроз на ранніх стадіях їх формування та поширення.

Запропонована система створює можливість не лише реагувати на вже поширені маніпуляції, але й прогнозувати напрямки майбутніх інформаційно-психологічних операцій противника. Це особливо важливо в умовах затяжної гібридної війни, коли когнітивний вплив є постійним і системним.

Таким чином, NCTDS може стати ключовим елементом національної системи когнітивної безпеки України, забезпечивши перехід від фрагментарних заходів протидії до єдиної проактивної, технологічно та психологічно обґрунтованої системи захисту суспільства від маніпулятивного впливу.

3.2. Методи протидії інформаційним кампаніям, націленим на когнітивні вразливості населення

Якщо технічні засоби детекції, дозволяють виявляти загрози на ранніх етапах їх формування та поширення, то безпосередня протидія інформаційно-психологічним операціям вимагає принципово іншого, комплексного та багаторівневого підходу. Сучасні технічні системи досить ефективно ідентифікують джерельну дезінформацію, проте демонструють суттєві обмеження при роботі з контентом, який формально не містить відвертої брехні, але спеціально сконструйований для експлуатації когнітивних викривлень, емоційного зараження, групової поляризації та підтверджувального упередження. Саме тому в умовах гібридної війни стратегічно необхідним стає перехід від переважно реактивних заходів (спростування вже поширеного контенту) до проактивних стратегій захисту, спрямованих на підвищення когнітивної стійкості населення ще до моменту зіткнення з маніпулятивним впливом.

Стратегічні цілі протидії когнітивним загрозам можна сформулювати таким чином:

- підвищення індивідуальної когнітивної стійкості громадян через розвиток критичного мислення, емоційної регуляції та навичок розпізнавання маніпуляцій;
- зміцнення колективної резистентності суспільства шляхом формування соціальних норм критичного сприйняття інформації та групової психологічної імунізації;
- обмеження впливу ворожих інформаційно-психологічних операцій на суспільні процеси, прийняття державних рішень та морально-психологічний стан населення;
- створення інтегрованої національної системи когнітивної безпеки, що забезпечує синергію технічних, психологічних, освітніх та інституційних інструментів захисту.

Досягнення цих цілей неможливе без ґрунтовного теоретичного підґрунтя. Теоретичні основи сучасних методів протидії формувалися на перетині когнітивної психології, соціальної психології та теорії комунікації.

Найважливішим парадигмовим зсувом останніх десятиліть стало визнання пріоритету проактивних стратегій над реактивними. Якщо раніше основна увага приділялася фактчекінгу та спростуванню дезінформації, то сьогодні дедалі більше дослідників схиляються до висновку, що превентивні заходи є значно ефективнішими у довгостроковій перспективі.

Серед ключових теоретичних моделей, які лежать в основі сучасних підходів до протидії, слід виділити три фундаментальні:

1. По-перше, теорію психологічної імунізації (McGuire, 1964), суттєво розвинену Roozenbeek, van der Linden та іншими дослідниками у 2019–2025 роках. Згідно з цією теорією, попереднє ознайомлення людини з ослабленою формою маніпулятивної техніки формує своєрідні «ментальні антитіла», що підвищують стійкість до повноцінної дезінформації, аналогічно до дії вакцини.
2. По-друге, модель ймовірності розробки (Elaboration Likelihood Model – ELM) Р. Петті та Дж. Какіоппо (1986), яка пояснює, що переконання, сформовані через центральний маршрут обробки інформації (глибоке осмислення та критичний аналіз), є значно стійкішими, ніж ті, що виникли через периферійний маршрут (емоційне сприйняття).
3. По-третє, теорію психологічної реактивності, яка обґрунтовує існування ефекту бумеранга при прямому та агресивному спростуванні інформації. Ця модель підкреслює необхідність уникати конфронтаційного стилю комунікації, оскільки він часто призводить до ще більшої закріпленості хибних переконань.

Водночас програми психологічної імунізації та пребанкінгу (попереджувальне роз’яснення маніпулятивних технік) забезпечують стійкіший захист, який зберігається навіть при зіткненні з новими, раніше невідомими наративами [37].

Особливо цінним є те, що превентивні методи працюють не тільки проти конкретного контенту, але й проти цілих класів маніпулятивних технік. Це робить їх універсальним інструментом у динамічному інформаційному середовищі гібридної війни, де противник постійно змінює тактику.

Ефективна протидія інформаційно-психологічним операціям, що експлуатують когнітивні викривлення, неможлива в рамках одного рівня впливу. Сучасні гібридні загрози вимагають системного, багаторівневого підходу, який охоплює індивідуальний, груповий (спільнотний) та інституційний рівні. Така архітектура забезпечує синергетичний ефект: зміцнення стійкості на мікро-рівні підтримується соціальними нормами на мезо-рівні та системною державною політикою на макро-рівні.

Для систематизації методів протидії пропонується сучасна класифікація, що враховує специфіку українського суспільства в умовах тривалої гібридної війни. Багаторівнева класифікація методів протидії когнітивним загрозам відображено у табл. 3.3.

Таблиця 3.3

Багаторівнева класифікація методів протидії когнітивним загрозам

Рівень застосування	**Основні методи протидії**	**Ключові завдання**	**Умови ефективності**	**Очікуваний ефект**
Індивідуальний	Психологічна імунізація, тренінги критичного мислення, цифрова гігієна, емоційна саморегуляція	Формування персональних навичок розпізнавання маніпуляцій та стійкості до когнітивних викривлень	Регулярність тренувань, гейміфікація, адаптація під вікову та соціальну групу	Індивідуальне посилення в довгостроковому плані
Груповий (спільнотний)	Спільнотні програми імунізації, розробка контрнаративів, групова рефлексія, робота з лідерами думок	Зміцнення стійкості соціальних спільнот та зниження групової поляризації	Висока довіра до джерела, емоційна залученість, системність заходів	Соціальне посилення в довгостроковому плані
Інституційний (державний)	Національні стратегії когнітивної безпеки, регулювання платформ, єдина комунікаційна політика, міжнародна співпраця	Створення системних умов захисту інформаційного простору	Міжвідомча координація, законодавче забезпечення, стабільне фінансування	Системне посилення держави в довгостроковому плані

1. Індивідуальний рівень є фундаментальним. Саме на цьому рівні відбувається первинне формування когнітивної стійкості. Основним інструментом виступає психологічна імунізація, яка дозволяє людині розпізнавати маніпулятивні техніки ще до їх масового поширення. Дослідження демонструють, що навіть короткі (10–15 хвилин) імунізаційні інтервенції значно підвищують здатність людини протистояти дезінформації.

Для України особливо актуальним є адаптування цих програм під різні вікові групи. Наприклад, для молоді ефективними є гейміфіковані формати, тоді як для старшого покоління — більш традиційні освітні модулі з акцентом на емоційну регуляцію та евристики.

2. Груповий (спільнотний) рівень дозволяє масштабувати індивідуальну стійкість через соціальні механізми. На цьому рівні ключовим є використання ефекту соціального доказу у позитивному ключі. Спільнотні програми імунізації в навчальних закладах, військових підрозділах, волонтерських організаціях та територіальних громадах сприяють формуванню колективних норм критичного сприйняття інформації. Групова рефлексія допомагає зменшити ефект ехокамер і групової поляризації, які активно експлуатуються російськими ІПСО.

3. Інституційний рівень виконує системоутворюючу функцію. Він передбачає розробку та впровадження державної політики когнітивної безпеки, законодавче регулювання діяльності великих цифрових платформ, створення єдиних комунікаційних стандартів для державних органів та координацію з міжнародними партнерами (NATO, ЄС, EUvsDisinfo). Без сильної інституційної підтримки окремі освітні та психологічні заходи залишаються фрагментарними та малоефективними.

Багаторівневий підхід є оптимальним саме для України з кількох причин:

– По-перше, українське суспільство характеризується високою горизонтальною довірою (довіра між людьми) при відносно низькій

вертикальній (довіра до інститутів). Це означає, що груповий рівень може стати потужним каталізатором змін.

- По-друге, тривала гібридна війна створює умови постійного інформаційного тиску, за яких реактивні методи швидко вичерпують себе.
- По-третє, лише поєднання всіх трьох рівнів дозволяє досягти необхідної глибини та масштабності захисту.

Таким чином, багаторівневий підхід не лише теоретично обґрунтований, але й практично необхідний для підвищення когнітивної стійкості українського суспільства в умовах сучасної гібридної війни. Його реалізація вимагає чітких практичних механізмів, які будуть розглянуті в наступному підрозділі.

Найперспективнішим практичним інструментом підвищення когнітивної стійкості населення в умовах гібридної війни є превентивні стратегії, серед яких центральне місце посідають психологічна імунізація та пребанкінг (попереджувальне роз'яснення маніпулятивних технік). На відміну від реактивних методів, ці підходи спрямовані не на спростування вже поширеного контенту, а на формування стійкості до маніпуляцій ще до їх масового поширення [38].

Для забезпечення ефективності психологічної імунізації необхідно адаптувати її інструменти до соціокультурної специфіки українського суспільства, особливостей сприйняття інформації різними демографічними групами та домінуючих каналів комунікації. Саме тому пропонується впровадити на системному рівні комплекс коротких інтерактивних імунізаційних модулів тривалістю 10–15 хвилин, орієнтованих на превентивне формування когнітивної стійкості.

1. Кожен модуль повинен бути присвячений конкретній маніпулятивній техніці або когнітивному викривленню (підтверджувальне упередження, емоційне зараження, ефект якоря, групова поляризація, евристика доступності тощо). Рекомендована структура модуля включає:

– Теоретичне пояснення механізму маніпулятивної техніки та відповідного когнітивного викривлення;

– Аналіз реальних прикладів з практики російських інформаційно-психологічних операцій проти України у 2022–2026 роках;

– «Ослаблену» версію маніпуляції для практичного тренування розпізнавання;

– Інтерактивні тестові завдання на ідентифікацію техніки та вироблення контрреакції.

2. Матеріали мають бути сегментовані відповідно до психологічних особливостей, рівня цифрової грамотності та домінуючих каналів споживання інформації цільових груп:

– школярі 9–11 класів;

– студенти вищих навчальних закладів;

– військовослужбовці;

– державні службовці;

– пенсіонери та старше покоління.

Для кожної групи передбачається адаптація за рівнем складності, мовним стилем (українська / російська), форматом подання (анімація, відеоролики, інфографіка, текст) та контекстними прикладами. Така диференціація дозволяє підвищити релевантність і ефективність імунізації.

3. Важливим елементом проактивної стратегії є пребанкінг — попереджувальна комунікація перед очікуваними інформаційними хвилями. Такі кампанії доцільно проводити напередодні потенційно критичних періодів: хвиль мобілізації, опалювального сезону, обговорення міжнародної допомоги, важливих політичних подій тощо [39].

Формат: короткі просвітницькі ролики, інфографіка та тематичні дописи в офіційних Telegram-каналах державних органів і популярних медіа. Ключове повідомлення має бути чітким: «Росія традиційно застосовує таку техніку — ось як її розпізнати».

4. Перспективним інструментом масової психологічної імунізації є створення національного гейміфікованого мобільного застосунку «Когнітивний щит». Застосунок може стати центральним елементом системної роботи з підвищення когнітивної стійкості населення. Основний функціонал включає:

- щоденні короткі тренувальні сесії;
- система рівнів, досягнень і винагород;
- персоналізовані рекомендації залежно від віку, інтересів і попередніх результатів користувача;
- інтеграція з Національною системою когнітивної детекції (NCTDS) для використання актуальних прикладів маніпуляцій;
- статистика індивідуального прогресу та порівняння з середніми показниками.

5. Гейміфікація суттєво підвищує залученість і довгострокову ефективність програм імунізації, особливо серед молоді.

Контрнаративи повинні будуватися не на прямому спростуванні ворожих тез (що часто провокує ефект бумеранга), а на пропозиції альтернативних когнітивних рамок, які відповідають українським цінностям, історичному досвіду та колективній ідентичності.

Приклади контрнаративів:

- Замість спростування тези «влада зраджує народ» — наратив «Українці об’єднані спільною метою та спільною відповідальністю».
- Замість спростування наративу «війна до останнього українця» — наратив «Ми воюємо за майбутнє наших дітей і свободу наступних поколінь».

Контрнаративи мають бути емоційно позитивними, базуватися на реальних історіях героїзму, стійкості, солідарності та взаємодопомоги. Їхнє поширення рекомендується здійснювати через мережу довірених джерел (офіційні медіа, відомі громадські діячі, військові блогери) [40].

Для системної протидії когнітивним загрозам в умовах тривалої гібридної війни недостатньо розрізнених заходів. Необхідне створення цілісної інтегрованої національної системи когнітивної безпеки, яка поєднує технічні,

психологічні, комунікаційні та правові інструменти в єдиний взаємопов'язаний механізм. Пропонується створити інтегровану модель державно-громадсько-приватного партнерства (ДГПП), що базується на чотирьох взаємодоповнюючих блоках і забезпечує синергетичний ефект між ними.

1. Технічний блок

Ядром даного блоку є Національна система когнітивної детекції (NCTDS). Система виконує функцію постійного моніторингу інформаційного простору, раннього виявлення маніпулятивного контенту та генерації актуальних даних для інших блоків системи.

Основні функції технічного блоку:

- Моніторинг у реальному часі інформаційного простору (Telegram, соціальні мережі, новинні ресурси);
- Мультимодальний аналіз контенту;
- Розрахунок індексу когнітивного маніпулятивного потенціалу (Cognitive Manipulation Score — CMS);
- Автоматичне сповіщення відповідальних органів та партнерів.

Технічний блок виконує роль «очей» усієї національної системи, забезпечуючи оперативність виявлення загроз і науково обґрунтовану основу для подальших психологічних та комунікаційних заходів.

2. Освітньо-психологічний блок

Цей блок відповідає за масове підвищення когнітивної стійкості населення та формування довгострокових навичок критичного мислення. Він включає:

- Впровадження системної психологічної імунізації в шкільних та університетських освітніх програмах;
- Розвиток та поширення національного гейміфікованого застосунку «Когнітивний щит»;
- Проведення спеціалізованих тренінгових програм для вразливих професійних і соціальних груп (військовослужбовці, державні службовці, журналісти, волонтери);

– Інституціоналізацію розвитку критичного мислення як обов’язкової компетентності в державних освітніх стандартах.

Головне завдання освітньо-психологічного блоку — перетворення пасивних споживачів інформації на активних, критично мислячих громадян, стійких до маніпулятивного впливу.

3. Комунікаційний блок відповідає за оперативну протидію інформаційним загрозам та формування альтернативних наративів у суспільній свідомості. Його основне завдання — трансформувати дані, отримані від Національної системи когнітивної детекції (NCTDS), у конкретні інформаційні заходи, що випереджають або нейтралізують поширення ворожих наративів.

Основні інструменти комунікаційного блоку включають:

– Систему швидкого пребанкінгу (попереджувальних кампаній), що запускаються при виявленні ознак підготовки інформаційної операції;
– Розробку та системне поширення ефективних контрнаративів, побудованих на українських цінностях та реальному досвіді;
– Формування єдиної комунікаційної політики державних органів та установ;
– Координацію зусиль з незалежними медіа, фактчекінговими організаціями та громадськими ініціативами.

Комунікаційний блок виконує роль «голоса» національної системи когнітивної безпеки, забезпечуючи оперативний перехід від виявлення загрози до проактивного інформаційного впливу.

4. Правовий та регуляторний блок

Правовий та регуляторний блок створює нормативно-правову основу функціонування всієї системи когнітивної безпеки. Він забезпечує баланс між ефективним протистоянням інформаційній агресії та дотриманням конституційних принципів свободи слова та права на інформацію.

До ключових напрямів діяльності блоку належать:

- Законодавче регулювання діяльності великих цифрових платформ, зокрема запровадження обов'язкової співпраці з NCTDS та надання необхідних метаданих;
- Встановлення юридичної відповідальності за скоординоване поширення маніпулятивного контенту та діяльність бот-мереж;
- Розробка механізмів захисту свободи слова при одночасному протидії інформаційній агресії іноземних держав;
- Розвиток міжнародної правової співпраці в рамках НАТО, Європейського Союзу та двосторонніх угод.

Усі чотири блоки функціонують у єдиному операційному контурі, що забезпечує системну синергію:

NCTDS (технічний блок) виявляє потенційну загрозу → дані автоматично передаються до комунікаційного та освітньо-психологічного блоків для запуску оперативного prebunking та цільової імунізації → правовий блок забезпечує нормативне підґрунтя, координацію та юридичний супровід заходів.

Така архітектура дозволяє здійснювати перехід від переважно реактивного режиму протидії до проактивного, превентивного підходу, суттєво підвищуючи загальну ефективність системи.

Оптимальною організаційною формою реалізації інтегрованої національної системи когнітивної безпеки визнається модель державно-громадсько-приватного партнерства (ДГПП). Вона передбачає чіткий розподіл ролей між ключовими стейкхолдерами:

- Держава забезпечує стратегічне керівництво, централізоване фінансування, доступ до державних даних, законодавчу базу та координацію між відомствами.
- Громадський сектор (фактчекінгові організації, медіа, громадські об'єднання, волонтери) надає легітимність, суспільну довіру, оперативність реакції та незалежну експертизу.

– Приватний сектор (телекомунікаційні оператори, технологічні платформи, ІТ-компанії) забезпечує технічні можливості, алгоритмічну інфраструктуру, канали поширення та необхідні дані.

Найперспективнішою є інтегрована стратегія, що поєднує технічні засоби детекції, превентивні психологічні програми, оперативні реактивні інструменти та системні державні заходи, які включають співпрацює між державними органами, громадським сектором, освітніми установами та технологічними компаніями в рамках моделі державно-громадсько-приватного партнерства (ДГПП). Лише такий комплексний підхід дозволяє суттєво підвищити когнітивну стійкість українського суспільства в умовах тривалої гібридної війни.

Висновки до розділу 3

Проаналізовано сучасні технічні та програмні засоби детекції маніпулятивного контенту. Встановлено, що існуючі рішення демонструють помірну ефективність при виявленні фактологічної дезінформації, але суттєво поступаються у розпізнаванні складних когнітивно орієнтованих маніпуляцій, мультимодального контенту та персоналізованих наративів.

Запропоновано концепцію інтегрованої національної системи когнітивної детекції (National Cognitive Threat Detection System — NCTDS). Розроблено її архітектуру, визначено чотири основні модулі (збору та моніторингу, мультимодального аналізу, психологічної оцінки, прийняття рішень та реагування), сформульовано ключові принципи функціонування (мультимодальність, гібридність, адаптивність, проактивність).

Запропоновано модель державно-громадсько-приватного партнерства (ДГПП), що поєднує технічний, освітньо-психологічний, комунікаційний та правовий блоки. Розроблено конкретні практичні інструменти: короткі імунізаційні модулі, гейміфікований застосунок «Когнітивний щит», систему пребанкінгу та механізми формування контрнаративів.

Отримані результати свідчать, що ефективний захист від когнітивних викривлень можливий лише за умови системного поєднання проактивної

технічної детекції та психологічно обґрунтованих превентивних заходів. Запропоновані в розділі розробки (NCTDS та модель ДГПП) становлять практичний внесок автора у посилення когнітивної безпеки України в умовах гібридної війни.

ВИСНОВКИ

Досліджено та розкрито концепцію двоїстої моделі мислення Д. Канемана (Система 1 і Система 2), природу евристик та когнітивних спотворень як системних особливостей людського мислення. Розроблено багатофакторну класифікацію когнітивних викривлень у контексті споживання контенту в кіберпросторі, яка виділяє три основні групи: когнітивні, емоційно-афективні та соціально-групові. Проаналізовано специфіку їх прояву залежно від типу контенту та впливу алгоритмічної персоналізації. Показано перехід індивідуальних когнітивних спотворень у колективні соціально-психологічні ефекти (емоційне зараження, соціальний доказ, групова поляризація), що суттєво впливає на суспільну стійкість, особливо в умовах гібридної війни.

Проведено аналіз механізмів використання когнітивних викривлень в інформаційно-психологічних операціях. Встановлено, що когнітивні викривлення є ключовим інструментом сучасних ІПСО Російської Федерації проти України. Проаналізовано еволюцію наративів, сегментацію цільових аудиторій та адаптацію маніпулятивних технік на різних етапах війни (2014–2025 рр.). Здійснено систематизацію маніпулятивних технік за типом експлуатованих викривлень та розглянуто ключові теоретичні моделі впливу. Оцінено ефективність методів маніпулятивного впливу за допомогою інтегрованого підходу. Доведено обмежену ефективність реактивних методів протидії (фактчекінг) та високу перспективність превентивних стратегій психологічної імунізації.

Розроблено практичні рекомендації щодо захисту від когнітивних загроз. Запропоновано концепцію та архітектуру Національної системи когнітивної детекції (NCTDS) — гібридного рішення, що поєднує мультимодальний ШІ-аналіз з психологічною оцінкою контенту. Обґрунтовано необхідність комплексного багаторівневого підходу до протидії (індивідуальний, груповий, інституційний) та запропоновано модель державно-громадсько-приватного

партнерства (ДГПП) як оптимальний механізм її реалізації. Особливу увагу приділено превентивним інструментам — психологічній імунізації, гейміфікації та пребанкінгу.

Отримані результати підтверджують, що ефективний захист від когнітивних викривлень у кіберпросторі можливий лише за умови системного поєднання теоретичного осмислення, сучасних технічних засобів детекції та проактивних психологічних стратегій підвищення когнітивної стійкості. Запропоновані в роботі концепції NCTDS та модель ДГПП вносять певний внесок у розвиток теорії та практики когнітивної безпеки України та можуть бути використані для посилення національної стійкості в умовах тривалої гібридної війни.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Newman N. et al. Digital News Report 2025. Reuters Institute for the Study of Journalism, University of Oxford, 2025. 180 с. URL: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2025-06/Digital_News-Report_2025.pdf.

2. The Global Risks Report 2025. 20th ed. Geneva : World Economic Forum, 2025. 122 с. URL: https://reports.weforum.org/docs/WEF_Global_Risks_Report_2025.pdf.

3. International Opinion on Global Threats. Pew Research Center, 2025. URL: https://www.pewresearch.org/wp-content/uploads/sites/20/2025/08/pg_2025.08.19_global-threats_reporta.pdf.

4. Zhou Y., Shen L. Processing of misinformation as motivational and cognitive biases. Frontiers in Psychology. 2024. Vol. 15. DOI: https://doi.org/10.3389/fpsyg.2024.1430953.

5. Tversky A., Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science. 1974. Vol. 185, No. 4157. P. 1124–1131. DOI: https://doi.org/10.1126/science.185.4157.1124.

6. Cognitive Psychology. Lumen Learning,URL: https://courses.lumenlearning.com/suny-hvcc-cogonitivepsychology/chapter/chapter-1/.

7. Tversky A., Kahneman D. Judgment under Uncertainty: Heuristics and Biases. URL: https://sites.socsci.uci.edu/~bskyrms/bio/readings/tversky_k_heuristics_biases.pdf.

8. Judgment under Uncertainty: Heuristics and Biases / ed. by D. Kahneman, P. Slovic, A. Tversky. Cambridge : Cambridge University Press, 1982. 555 с. DOI: https://doi.org/10.1017/CBO9780511809477.

9. Kahneman D. Thinking, Fast and Slow. New York : Farrar, Straus and Giroux, 2011. 499 с. URL:

https://dn790002.ca.archive.org/0/items/DanielKahnemanThinkingFastAndSlow/Daniel%20Kahneman-Thinking%2C%20Fast%20and%20Slow%20%20.pdf.

10. Kahneman D. Thinking, Fast and Slow. New York : Farrar, Straus and Giroux, 2011. 499 c. URL: https://dn790002.ca.archive.org/0/items/DanielKahnemanThinkingFastAndSlow/Daniel%20Kahneman-Thinking%2C%20Fast%20and%20Slow%20%20.pdf.

11. Когнітивна війна: людська свідомість як поле бою. Ужгородський національний університет, [2024]. URL: https://dspace.uzhnu.edu.ua/bitstreams/851c77ed-f22b-4478-b5bd-55fcda1fc7f8/download.

12. Cognitive Biases in Digital Decision-Making: How Consumers Navigate Information Overload. ACR Journal. [2025]. URL: https://www.researchgate.net/publication/384638015_The_Impact_of_Cognitive_Biases_on_Consumer_Decision-Making.

13. Вплив когнітивних викривлень на розвиток економіки. Менеджмент. 2023. С. 287–292. URL: https://science.lpnu.ua/sites/default/files/journal-paper/2023/sep/31287/menedzhment-287-292.pdf.

14. Identify Biases in the Digital World for a More Open and Aware Mind. TSW.URL: https://www.tsw.it/en/journal-eng/research-experiences/identify-biases-in-the-digital-world-for-a-more-open-and-aware-mind/.

15. The psychological drivers of misinformation belief and its resistance to correction. Nature Reviews Psychology. 2022. DOI: https://doi.org/10.1038/s44159-021-00006-y.

16. Індекс медіаграмотності українців залишається високим, але за рік знизився рівень цифрової компетентності: дослідження «Детектора медіа». Detector Media. 2025. 06 трав. URL: https://ms.detector.media/mediadoslidzhennya/post/37877/2025-05-06-indeks-mediagramotnosti-ukraintsiv-zalyshaietsya-vysokym-ale-za-rik-znyzyvsya-riven-tsyfrovoi-kompetentnosti-doslidzhennya-detektora-media/.

17. Оцінка змін, що відбулися в українському суспільстві за час повномасштабної війни (листопад 2025 р.). Київ : Центр Разумкова, 2025. URL: https://razumkov.org.ua/napriamky/sotsiologichni-doslidzhennia/otsinka-zmin-shcho-vidbulysia-v-ukrainskomu-suspilstvi-za-chas-povnomasshtabnoi-viiny-lystopad-2025r.

18. Cognitive Attacks in Russian Hybrid Warfare. Institute for Security and International Studies. URL: https://isij.eu/article/cognitive-attacks-russian-hybrid-warfare.

19. Cyber Influence Defense: Applying the DISARM Framework to a Cognitive Hacking Case from the Romanian Digital Space. ResearchGate. 2024. URL: https://www.researchgate.net/publication/382169419_Cyber_influence_defense_Applying_the_DISARM_framework_to_a_cognitive_hacking_case_from_the_Romanian_digital_space.

20. Modern Cognitiv Modern Cognitive Operations and Hybrid W ations and Hybrid Warfare. Journal of Strategic Security. URL: https://digitalcommons.usf.edu/cgi/viewcontent.cgi?article=2032&context=jss.

21. The Cognitive Battlefield of Hybrid Warfare. NATO Defense College Foundation, 2025. 7 листопада. URL: https://www.natofoundation.org/food/the-cognitive-battlefield-of-hybrid-warfare/.

22. Claverie B., du Cluzel F. The Cognitive Warfare Concept. NATO Innovation Hub, 2022. URL: https://innovationhub-act.org/wp-content/uploads/2023/12/CW-article-Claverie-du-Cluzel-final_0.pdf.

23. Thomas T. L. Russia's Reflexive Control Theory and the Military. The Journal of Slavic Military Studies. 2004. Vol. 17, No. 2. P. 237–256. DOI: https://doi.org/10.1080/13518040490450529.

24. Lazer D. M. J. et al. The science of fake news. Science. 2018. Vol. 359, Issue 6380. P. 1094–1096. DOI: https://doi.org/10.1126/science.aao2998.

25. 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.

26. The science of effective learning with a focus on spacing and retrieval practice .URL: https://www.researchgate.net/publication/362093173_The_science_of_effective_learning_with_a_focus_on_spacing_and_retrieval_practice

27. Chan M. S. et al. A meta-analysis of correction effects in science-relevant misinformation. [Journal], 2023. URL: https://socialactionlab.org/wp-content/uploads/2024/01/Chan_A-meta-analysis-of-correction-effects-in-science-relevant-misinformation_2023.pdf.

28. Walter N., Murphy S. T. A meta-analytic approach to correction of misinformation. Communication Monographs. 2018. DOI: https://doi.org/10.1080/03637751.2018.1467564.

29. Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News. Journal of Cognition. 2020. DOI: https://doi.org/10.5334/joc.91. URL: https://journalofcognition.org/articles/10.5334/joc.91.

30. Zhou Y., Shen L. Processing of misinformation as motivational and cognitive biases. Frontiers in Psychology. 2024. Vol. 15. DOI: https://doi.org/10.3389/fpsyg.2024.1430953.

31. The Role of Trust and Attitudes toward Democracy in the Dissemination of Disinformation—a Comparative Analysis of Six Democracies. Digital Journalism. 2023. URL: https://www.tandfonline.com/doi/full/10.1080/21670811.2023.2200196.

32. Next-level partnership: bolstering EU-NATO cooperation to counter hybrid threats. EU Institute for Security Studies. URL: https://www.iss.europa.eu/publications/briefs/next-level-partnership-bolstering-eu-nato-cooperation-counter-hybrid-threats

33. Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. PLOS ONE. 2017. DOI: https://doi.org/10.1371/journal.pone.0175799.

34. Psychological inoculation improves resilience to and reduces willingness to share vaccine misinformation.Scientific Reports. 2025. DOI: https://doi.org/10.1038/s41598-025-09462-5.

35. Current Opinion in Behavioral Sciences. DOI: https://www.sciencedirect.com/journal/current-opinion-in-behavioral-sciences

36. The Elaboration Likelihood Model of Persuasion Advances in Experimental Social Psychology. DOI: https://doi.org/10.1016/S0065-2601(08)60214-2

37. Reflective and Impulsive Determinants of Social Behavior. Psychological Science in the Public Interest. 2004. Vol. 5, No. 3. DOI: https://doi.org/10.1207/s15327957pspr0803_1

38. Psychological booster shots targeting memory increase long-term resistance against misinformation. Nature Communications. 2025. DOI: https://doi.org/10.1038/s41467-025-57205-x.

39. Prebunking misinformation techniques in social media feeds: Results from an Instagram field study. The Harvard Kennedy School Misinformation Review. URL: https://misinforeview.hks.harvard.edu/article/prebunking-misinformation-techniques-in-social-media-feeds-results-from-an-instagram-field-study/.

40. Preventive Strategies Against Disinformation: A Study on Digital and Information Literacy Activities Led by Fact-Checking Organisations. PMC / Journal. 2024. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC12457897/.