

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

Факультет соціології і права

Кафедра соціології

«На правах рукопису»
УДК _____

ДО ЗАХИСТУ ДОПУЩЕНО
Завідувач кафедри
_____ Максим ЄНІН
«____» _____ 2026 р.

Магістерська дисертація

на здобуття ступеня магістра

за освітньо-науковою програмою «Аналітика соціальних даних»

зі спеціальності 054 Соціологія

**на тему: «Трансформація ціннісних орієнтацій у дискурсі української
художньої літератури (1991–2025 рр.)»**

Виконав (-ла):

студент (-ка) II курсу, групи СЛ-41мн

Маїлян Армен Самвелович _____

Науковий керівник:

Доцент, кандидат соціологічних наук, доцент

Коржов Геннадій Олександрович _____

Рецензент:

Доцент кафедри історії, кандидат історичних наук, доцент,

Кізлова Антоніна Анатоліївна _____

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних посилань.
Студент (-ка) _____

Київ – 2026 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Факультет соціології і права
Кафедра соціології

Рівень вищої освіти – другий (магістерський)

Спеціальність – 054 Соціологія

Освітньо-наукова програма «**Аналітика соціальних даних**»

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Максим ЄНІН
« ____ » _____ 2026 р.

ЗАВДАННЯ
на магістерську дисертацію студенту

Маїлян Армен Самвелович

1. Тема дисертації «Трансформація ціннісних орієнтацій у дискурсі української художньої літератури (1991–2025 рр.)», науковий керівник дисертації Коржов Геннадій Олександрович, кандидат соціологічних наук, доцент, затверджені наказом по університету від «30» березня 2026р. №1318-с.

2. Термін подання студентом дисертації: 15.05.2026

3. Об'єкт дослідження: репрезентації ціннісних орієнтацій у дискурсі україномовної художньої прози 1991–2025 рр.

4. Предмет дослідження: закономірності динаміки та жанрової стратифікації ціннісних репрезентацій у чотирьох хронологічних зрізах (1991–2003, 2004–2013, 2014–2021, 2022–2025 рр.), виявлені засобами інтегрованого текстоцентричного аналізу із застосуванням автоматизованого аксіологічного кодування, жанрової стандартизації, валідаційних процедур та діагностичного зіставлення з даними соціологічних опитувань.

5. Перелік завдань, які потрібно розробити:

- уточнити соціологічне розуміння цінностей, ціннісних орієнтацій, ціннісних систем і художнього дискурсу як джерела дослідження ціннісних

репрезентацій та обґрунтувати текстоцентричний підхід як методологічну рамку такого аналізу;

- обґрунтувати соціально-історичну рамку аналізу українського суспільства 1991–2025 рр. та визначити хронологічні зрізи, у межах яких досліджується художній дискурс;
- сформувати корпус україномовної художньої прози 1991–2025 рр., визначити критерії відбору творів, жанрову класифікацію, одиниці аналізу та процедури підготовки аналітичного датасету;
- розробити й перевірити інструментарій автоматизованого аксіологічного контент-аналізу, включаючи категоріальну систему, правила кодування, процедури валідації та визначення методологічних меж підходу;
- проаналізувати та інтерпретувати динаміку репрезентацій ціннісних категорій у хронологічних зрізах із застосуванням жанрової стандартизації, перевірки теоретичних очікувань та якісного аналізу показових і контрастних кейсів;
- виявити та проінтерпретувати співвідношення між корпусними показниками активності ціннісних маркерів і даними соціологічних опитувань (WVS, EVS, національні моніторинги) як двома незалежними оптиками фіксації ціннісних репрезентацій.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу: всього в праці використано 11 рисунків та 16 таблиць.

7. Орієнтовний перелік публікацій:

- Маїлян, А. С. (2026). Методи аналізу ціннісних систем у художньому дискурсі: текстоцентричний підхід. Вісник науки та освіти. Серія: Філологія. [https://doi.org/10.52058/2786-6165-2026-4\(46\)-1444-1491](https://doi.org/10.52058/2786-6165-2026-4(46)-1444-1491).
- Mailian, A., & Korzhov, H. (2026, 16 квітня). Analysis of value narratives in fiction as a factor of social influence in the digital media environment. У Фундаментальні та прикладні проблеми суспільства: історія, сьогодення, майбутнє: Матеріали III Міжнародної науково-практичної онлайн-конференції (С. 57–60). Державний торговельно-економічний університет. <https://doi.org/10.31617/k.knute.2026-04-16>.
- Маїлян, А. С. (2025, 27–28 лютого). Комплексне застосування кількісних і якісних методів аналізу тексту в ЗМІ для визначення можливості впливу на настрої, емоційний стан та політичну спрямованість читачів. У Україна у просторі соціальних комунікацій: суспільство, медіа, рефлексії війни: Матеріали конференції (С. 672–675). Державний торговельно-економічний університет. <https://doi.org/10.31617/k.knute.2025-02-27-28>.

8. Дата видачі завдання: вересень 2024

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів роботи	Примітка
	Визначення напрямку дослідження	Жовтень 2024	
	Формування джерельної бази дослідження	Листопад 2024 – січень 2025	
	Складання розгорнутого плану	Лютий - березень 2025	
	Написання вступу до роботи (визначення об'єкту, предмету, мети та завдань)	Квітень-червень 2025	
	Підготовка і написання I розділу дипломної роботи (Теоретико-методологічні засади дослідження трансформації ціннісних орієнтацій у дискурсі української художньої літератури)	Вересень-листопад 2025	
	Написання II розділу дипломної роботи (Корпус і інструментарій)	Грудень 2025 – лютий 2026	
	Написання III розділу (Валідація автоматизованого аксіологічного кодування).	Березень-квітень 2026	
	Написання IV розділу (Трансформація ціннісних орієнтацій у художньому дискурсі 1991–2025 рр.), висновків та коригування списку літератури.	Травень 2026	
	Передача магістерської дисертації на кафедру	15 травня 2026	

Студент

Армен МАІЛЯН

Науковий керівник

Геннадій КОРЖОВ

АНОТАЦІЯ

Маїлян А. С. Трансформація ціннісних орієнтацій у дискурсі української художньої літератури (1991–2025 рр.). — На правах рукопису.

Магістерська дисертація за спеціальністю 054 Соціологія. — Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», кафедра соціології. — Київ, 2026. — **125 с., 11 іл., 16 табл.**, список джерел з **158 найменувань, 16 додатків.**

Магістерську дисертацію присвячено соціологічній реконструкції трансформації ціннісних орієнтацій, репрезентованих у дискурсі україномовної художньої прози 1991–2025 рр. У роботі розроблено й апробовано інтегрований текстоцентричний підхід, що поєднує автоматизований аксіологічний контент-аналіз із процедурами валідації, жанровою стандартизацією та якісною інтерпретацією показових кейсів. Сформовано корпус обсягом 664 твори 235 авторів і 1 375 333 абзаци за чотирма хронологічними зрізами та п'ятьма жанровими блоками; розроблено категоріальну систему із семи ціннісних категорій, що поєднує чотири агреговані виміри моделі Ш. Шварца з трьома контекстно-специфічними категоріями українського художнього дискурсу — національною ідентичністю, травмою війни й історичних катастроф, свободою і колективною суб'єктністю.

Встановлено статистично значущий спадний тренд для всіх семи категорій при малих розмірах ефекту; жанрова стандартизація істотно послаблює тренди для п'яти категорій і зберігає переважно внутрішньожанровий характер спаду для самотрансцендентності та самоствердження. Виявлено емпіричний феномен дисоціації режимів репрезентації цінностей: значна частина ціннісних орієнтацій, що в ранніх жанрових сегментах корпусу прямо лексично маркувалася, у пізніших текстах переходить у фонові, наративні або дискурсивно опосередковані форми. Діагностичне зіставлення з даними WVS, EVS і

національних опитувань зафіксувало розбіжність між опитувальною і корпусною оптиками щодо більшості категорій, інтерпретовану як індикатор зміни режиму репрезентації, а не безпосередньо самих ціннісних орієнтацій.

Ключові слова: *ціннісні орієнтації, ціннісна трансформація, художній дискурс, українська художня проза, текстоцентричний підхід, автоматизований аксіологічний контент-аналіз, жанрова стандартизація, теорія цінностей Ш. Шварца, корпусний аналіз, постсоціалістичне суспільство.*

SUMMARY

Mailian A. S. Transformation of Value Orientations in the Discourse of Ukrainian Fiction (1991–2025). — Manuscript.

Master's thesis on speciality C5 Sociology. – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Department of Sociology. – Kyiv, 2026. – 125 p., 11 fig., 16 tabl., 158 sources, 16 appendices.

The thesis is devoted to the sociological reconstruction of the transformation of value orientations represented in the discourse of Ukrainian-language fiction prose of 1991–2025. An integrated text-centred approach is developed and tested that combines automated axiological content analysis with validation procedures, genre standardization and qualitative interpretation of illustrative cases. A corpus of 664 works by 235 authors and 1,375,333 paragraphs was compiled across four chronological cross-sections and five genre blocks; a categorical system of seven value categories was developed, combining four aggregated dimensions of S. Schwartz's model with three context-specific categories of Ukrainian fictional discourse — national identity, war and historical catastrophe trauma, and freedom with collective subjectivity.

A statistically significant downward trend with small effect sizes was established for all seven categories; genre standardization substantially weakens the trends for five categories while preserving a predominantly intra-genre character of decline for self-transcendence and self-enhancement. An empirical phenomenon of dissociation of value representation regimes was revealed: a substantial part of value orientations that were directly lexically marked in earlier genre segments of the corpus shifts in later texts towards background, narrative or discursively mediated forms. Diagnostic juxtaposition with WVS, EVS and national survey data revealed a divergence between the survey-based and corpus-based optics for most categories, interpreted as an

indicator of a change in representation regime rather than of value orientations themselves.

Key words: *value orientations, value transformation, fictional discourse, Ukrainian fiction prose, text-centred approach, automated axiological content analysis, genre standardization, S. Schwartz value theory, corpus analysis, post-socialist society.*

ЗМІСТ

ВСТУП	1
РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ ТРАНСФОРМАЦІЇ ЦІННІСНИХ ОРІЄНТАЦІЙ У ДИСКУРСІ УКРАЇНСЬКОЇ ХУДОЖНЬОЇ ЛІТЕРАТУРИ	9
1.1. Цінності, ціннісні орієнтації та ціннісні системи: понятійний апарат, теоретичні моделі і соціологічна концептуалізація.....	9
1.2. Соціальна трансформація і ціннісні зрушення в постсоціалістичному суспільстві: теоретичний вимір та український контекст 1991–2025 рр.	18
1.3. Художній текст і художній дискурс як соціологічне джерело дослідження цінностей	28
1.4. Текстоцентричний підхід до аналізу ціннісних орієнтацій у художньому дискурсі.....	37
1.5. Теоретично мотивовані очікування про динаміку ціннісних репрезентацій.....	44
1.6. Висновки до розділу 1	46
РОЗДІЛ 2. КОРПУС І ІНСТРУМЕНТАРІЙ	49
2.1. Формування та підготовка корпусу текстів дослідження.....	49
2.2. Інструментарій автоматизованого аксіологічного контент-аналізу.....	61
2.3. Висновки до розділу 2	75
РОЗДІЛ 3. ВАЛІДАЦІЯ АВТОМАТИЗОВАНОГО АКСІОЛОГІЧНОГО КОДУВАННЯ.....	77
3.1. Внутрішня надійність кодування: α Криппендорфа	77
3.2. Зовнішня валідність: F1 проти трьох gold standards.....	82

3.3. Висновки до розділу 3	87
РОЗДІЛ 4. ТРАНСФОРМАЦІЯ ЦІННІСНИХ ОРІЄНТАЦІЙ У ХУДОЖНЬОМУ ДИСКУРСІ 1991–2025 РР.....	89
4.1. Первинна динаміка ціннісних категорій у корпусних даних	90
4.2. Жанрова стандартизація як корекція корпусних зміщень.....	96
4.3. Перевірка теоретичних гіпотез у світлі емпіричних результатів .	104
4.4. Якісна інтерпретація показових і контрастних кейсів	109
4.5. Зіставлення з даними WVS, EVS і національними опитуваннями: діагностична перевірка	113
4.6. Висновки до розділу 4	117
ВИСНОВКИ.....	121
РЕКОМЕНДАЦІЇ	124
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	126
ДОДАТКИ.....	142
Додаток А. Концепції цінностей в соціології	142
Додаток Б. Поняття та індикатори	143
Додаток В. Жанрова класифікація	145
Додаток Г. Демонстративний приклад JSON структури твору для обробки.....	148
Додаток Д. Кількісні показники робочого корпусу української художньої літератури 1991-2025	149
Додаток Е. Приклад corpus_stats.csv	153
Додаток Ж. Структура полів corpus_flat.csv	154
Додаток И. Кодбук дослідження	155
Додаток К. Процедура ручного кодування абзаців	169

Додаток Л. Попарна α -матриця узгодження кодувальників	174
Додаток М. Метрики precision / recall / F1	175
Додаток Н. Попарне співпадіння трьох gold standards між собою	176
Додаток П. Розширений якісний аналіз кейсів	177
Додаток Р. Попарні порівняння часток активних абзаців між хронологічними зрізами	180
Додаток С. Рисунки декомпозиції та Z-статистик.....	181
Додаток Т. Триангуляція корпусних результатів з даними World Values Survey і European Values Study. Опорні точки національних опитувань.	182

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

У переліку наведено умовні позначення, аббревіатури і скорочення, що використовуються в магістерській дисертації. Розширені дефініції наведено в основному тексті за місцем введення відповідного терміну.

α	коефіцієнт α Криппендорфа; міра міжкодувальної надійності для номінальних даних. Шкала інтерпретації, прийнята в роботі: $\geq 0,80$ — високе узгодження; $0,67-0,80$ — прийнятне; $0,40-0,67$ — слабке; $< 0,40$ — фактично відсутнє
A165, Y001, Y003, Y010, Y024, E035, E069_02, E117, F063, G006	кодові позначення окремих змінних WVS / EVS, використовувані у Розділі 4 і Додатку Т (повна семантика кожної змінної подана в Додатку Т)
API	Application Programming Interface; програмний інтерфейс взаємодії із зовнішнім сервісом (у роботі — звернення до моделей LLM)
CA-test	Cochran–Armitage trend test; тест на наявність монотонного тренду упорядкованих часток
Cohen’s h	стандартизований розмір ефекту для порівняння двох пропорцій
CON	Conservation; збереження (агрегована категорія за моделлю Ш. Шварца: безпека + конформність + традиція)

CONVERGENT	/ типи зіставлення корпусного тренду з опитувальними
DIVERGENT	/ оптиками (синхронний рух / розходження напрямків /
FLAT	синхронна стабільність)
Cramer's V	міра асоціації для таблиць спряженості
CSV	Comma-Separated Values; табличний текстовий формат (corpus_stats.csv, corpus_flat.csv)
DIF	Фонд «Демократичні ініціативи» імені Ілька Кучеріва (Democratic Initiatives Foundation)
ENH	Self-Enhancement; самоствердження (агрегована категорія за моделлю Ш. Шварца: досягнення + влада)
EVS	European Values Study; Європейське дослідження цінностей
Expanded-gold	об'єднаний консенсус усіх восьми кодувальників (5 людей + 3 LLM) із порогом $\geq 4/8$
F1	гармонічне середнє точності (precision) і повноти (recall)
FN	False Negative; хибно негативний результат класифікації
FP	False Positive; хибно позитивний результат класифікації
GPU	Graphics Processing Unit; графічний (тензорний) процесор для прискорення обчислень
H1, H2, H3	теоретично мотивовані очікування щодо динаміки UA- IDN, ENH і TRA відповідно (підрозділ 1.5)
H1s, H2s, H3s	жанрово-стратифіковані варіанти H1–H3

H4	методологічне очікування щодо впливу жанрової композиції корпусу на сирі тренди
Human-gold	мажоритарне голосування п'яти кодувальників-людей із порогом $\geq 3/5$
JSON	JavaScript Object Notation; формат структурованих даних, у якому представлено твори корпусу (Додаток Г)
LLM	Large Language Model; велика мовна модель (у роботі — Claude Opus 4.7, Gemini 2.5 Pro, GPT-4o)
LLM-gold	мажоритарне голосування трьох моделей LLM із порогом $\geq 2/3$
NLP	Natural Language Processing; обробка природної мови
OCR	Optical Character Recognition; оптичне розпізнавання символів
OPN	Openness to Change; відкритість до змін (агрегована категорія за моделлю Ш. Шварца: самостійність + стимуляція + гедонізм)
p-value	рівень статистичної значущості
PVQ-RR	Portrait Values Questionnaire — Revised (Refined); портретний опитувальник цінностей Ш. Шварца у переглянутій версії на 19 типів
SD	standard deviation; стандартне відхилення
SVS	Schwartz Value Survey; опитувальник цінностей Шварца
TN	True Negative; істинно негативний результат класифікації

TP	True Positive; істинно позитивний результат класифікації
TRA	Self-Transcendence; самотрансцендентність (агрегована категорія за моделлю Ш. Шварца: доброзичливість + універсалізм). У межах цієї роботи TRA не означає Tradition; традиція як базова цінність Ш. Шварца входить до CON
UA-FRE	свобода і колективна суб'єктність (контекстно-ціннісна категорія групи В)
UA-IDN	українська національна ідентичність і культурна самобутність (контекстно-ціннісна категорія групи В)
UA-WAR	травма війни, переміщення і пам'ять про катастрофи (контекстно-ціннісна категорія групи В)
UD	Universal Dependencies; міжнародний проєкт уніфікованої морфологічно-синтаксичної анотації корпусів
UPOS	Universal Part-of-Speech; UD-сумісна розмітка частин мови
WVS	World Values Survey; Всесвітнє дослідження цінностей
Z	Z-статистика тренд-тесту Cochran–Armitage
ДА	дискурс-аналіз
КДА	критичний дискурс-аналіз
КМІС	Київський міжнародний інститут соціології
п.п.	процентний пункт (одиниця різниці двох відсоткових значень; не плутати з відсотком зміни)

ВСТУП

Актуальність теми. Період 1991–2025 рр. охоплює складну й нелінійну траєкторію ціннісної трансформації українського суспільства: відновлення державної незалежності, пострадянську кризу 1990-х, хвилі громадянської мобілізації 2004 і 2013–2014 рр., російську агресію з 2014 р. та повномасштабне вторгнення 2022 р. Ці події змінювали не лише інституційний устрій, а й ієрархію суспільно значущих цінностей — безпеки, свободи, гідності, національної ідентичності, солідарності та справедливості. Соціологічна реконструкція цих зрушень важлива для розуміння траєкторії українського суспільства і культурних механізмів осмислення соціальних змін.

Класичні інструменти вимірювання цінностей (World Values Survey, European Values Study, Schwartz Value Survey, PVQ-RR) фіксують декларовані пріоритети, проте обмежені в діахронній реконструкції латентних, символічних і культурно опосередкованих форм цінностей у завершених історичних періодах, особливо у суспільствах, де ціннісні зміни були пов'язані з кризами, мобілізаціями та війною.

У зв'язку з цим зростає значущість текстоцентричних підходів, які працюють з текстовими артефактами відповідних історичних періодів. Україномовна художня проза 1991–2025 рр. є цінним джерелом такого аналізу, оскільки фіксує ціннісні репрезентації у формі, введеній у публічний культурний обіг, і дозволяє простежити зміну способів символічного оформлення цінностей у соціально-історично обґрунтованих періодах. Художній дискурс не є прямим відображенням суспільної свідомості, але репрезентує моральні опозиції, моделі належного, символічні ієрархії та культурні коди, важкодоступні для реконструкції лише опитувальними методами.

Окремою методологічною проблемою є поєднання кількісних і якісних процедур аналізу цінностей у художньому тексті: контент-аналіз дає змогу систематично фіксувати ціннісні маркери у великому корпусі художніх текстів, а дискурс-аналіз, наративний аналіз і комп'ютерно-асистовані процедури забезпечують додаткові рівні інтерпретації. Саме тому в роботі застосовано інтегрований текстоцентричний підхід, у межах якого автоматизоване кодування поєднується з валідацією, жанровою стандартизацією та якісним аналізом показових кейсів.

Додатковим аспектом актуальності є дефіцит методологічно прозорих процедур соціологічного аналізу ціннісних репрезентацій у великих масивах україномовних художніх текстів — зокрема, процедур коректного переходу від зафіксованих ціннісних маркерів до соціологічної інтерпретації ціннісних зрушень без ототожнення частотності слів із реальними цінностями авторів, читачів або суспільства. Розроблення такої процедури має академічне значення для соціології культури, аналізу соціальних даних і методології контент-аналізу, а також прикладне значення для моніторингу культурних трансформацій, гуманітарної політики, освітніх і медіастратегій.

Отже, актуальність теми визначається потребою соціологічного осмислення ціннісних трансформацій українського суспільства 1991–2025 рр., обмеженістю суто опитувальних методів для їхньої ретроспективної реконструкції та необхідністю розроблення україномовного інструментарію автоматизованого аксіологічного аналізу художнього дискурсу.

Зв'язок дослідження з науковими програмами, планами, темами. Роботу виконано в рамках ініціативної науково-дослідницької роботи кафедри соціології КПІ ім. Ігоря Сікорського «Інноваційні підходи до врегулювання сучасних соціально-політичних конфліктів» (державний реєстраційний номер: 0126U001993).

Мета дослідження — виявити та інтерпретувати трансформацію ціннісних орієнтацій, репрезентованих у дискурсі україномовної художньої прози 1991–2025 рр., шляхом розроблення й апробації інтегрованого текстоцентричного підходу, що поєднує автоматизоване кодування корпусу художніх текстів, процедури валідації, жанрову стандартизацію, діагностичне зіставлення з соціологічними джерелами та якісну інтерпретацію показових текстових кейсів.

Для досягнення поставленої мети у дослідженні сформульовано такі завдання:

- уточнити соціологічне розуміння цінностей, ціннісних орієнтацій, ціннісних систем і художнього дискурсу як джерела дослідження ціннісних репрезентацій та обґрунтувати текстоцентричний підхід як методологічну рамку такого аналізу;
- обґрунтувати соціально-історичну рамку аналізу українського суспільства 1991–2025 рр. та визначити хронологічні зрізи, у межах яких досліджується художній дискурс;
- сформувати корпус україномовної художньої прози 1991–2025 рр., визначити критерії відбору творів, жанрову класифікацію, одиниці аналізу та процедури підготовки аналітичного датасету;
- розробити й перевірити інструментарій автоматизованого аксіологічного контент-аналізу, включаючи категоріальну систему, правила кодування, процедури валідації та визначення методологічних меж підходу;
- проаналізувати та інтерпретувати динаміку репрезентацій ціннісних категорій у хронологічних зрізах із застосуванням жанрової стандартизації, перевірки теоретичних очікувань та якісного аналізу показових і контрастних кейсів;
- виявити та проінтерпретувати співвідношення між корпусними показниками активності ціннісних маркерів і даними соціологічних

опитувань (WVS, EVS, національні моніторинги) як двома незалежними оптиками фіксації ціннісних репрезентацій.

Об'єкт і предмет дослідження:

Об'єкт дослідження — репрезентації ціннісних орієнтацій у дискурсі україномовної художньої прози 1991–2025 рр.

Предмет дослідження — закономірності динаміки та жанрової стратифікації ціннісних репрезентацій у чотирьох хронологічних зрізах (1991–2003, 2004–2013, 2014–2021, 2022–2025 рр.), виявлені засобами інтегрованого текстоцентричного аналізу із застосуванням автоматизованого аксіологічного кодування, жанрової стандартизації, валідаційних процедур та діагностичного зіставлення з даними соціологічних опитувань.

Методи дослідження. У магістерській дисертації використано комплекс загальнонаукових, соціологічних, статистичних і комп'ютерно-асистованих методів. Теоретико-методологічна частина роботи ґрунтується на аналізі, синтезі, порівнянні, узагальненні та концептуалізації підходів до вивчення цінностей, ціннісних орієнтацій, соціальної трансформації, художнього тексту й художнього дискурсу.

Емпірична частина побудована на масштабному автоматизованому аксіологічному контент-аналізі україномовних художніх прозових текстів 1991–2025 рр. з опорою на категоріальну систему ціннісних маркерів; зіставність даних забезпечено процедурами мовної нормалізації та агрегування показників за хронологічними зрізами і жанровими блоками. Якість автоматизованого кодування перевірено за коефіцієнтом α Криппендорфа та показниками точності, повноти і F1-міри у зіставленні з референтними розмітками. Якісну інтерпретацію результатів виконано із залученням елементів критичного дискурс-аналізу та наративного аналізу.

Наукова новизна одержаних результатів полягає в розробленні й апробації інтегрованого текстоцентричного підходу до аналізу трансформації ціннісних орієнтацій, репрезентованих у дискурсі україномовної художньої прози 1991–2025 рр.

У роботі розроблено й апробовано процедуру автоматизованого аксіологічного контент-аналізу великого корпусу україномовної художньої прози, що поєднує чотири агреговані категорії моделі базових цінностей Ш. Шварца з трьома категоріями, релевантними українському соціокультурному контексту: національною ідентичністю та культурною самобутністю; травмою війни і пам'яттю про катастрофи; свободою і колективною суб'єктністю.

Удосконалено методику аналізу ціннісних репрезентацій у художньому дискурсі через поєднання автоматизованого кодування, процедур валідації, жанрової стандартизації та якісної інтерпретації текстових кейсів; запропоновано жанрово-чутливу процедуру інтерпретації корпусних трендів, яка відокремлює динаміку словниково зафіксованих маркерів від ефектів жанрової композиції корпусу.

Дістала подальшого розвитку методика перевірки автоматизованого аксіологічного кодування через зіставлення результатів алгоритму з Human-gold, LLM-gold та Expanded-gold, що дозволило виявити відмінності між людською, LLM-орієнтованою та словниковою інтерпретацією ціннісних маркерів у художньому тексті.

Емпірично окреслено й запропоновано для подальшого соціологічного обговорення робоче поняття дисоціації режимів репрезентації цінностей — розходження між декларативно-самозвітними індикаторами ціннісних орієнтацій і лексично-маркованою активністю відповідних ціннісних категорій у художньому дискурсі. У межах цієї роботи цей феномен інтерпретується не як

пряме спростування опитувальних або корпусних даних, а як показник різних способів соціокультурної фіксації цінностей.

Практичне значення одержаних результатів полягає в можливості використання сформованого корпусу, категоріальної системи, кодбуку, процедур автоматизованого кодування та підходів до валідації в подальших соціологічних, культурологічних, філологічних і міждисциплінарних дослідженнях українського художнього дискурсу.

Запропонований підхід може бути застосований у навчальних курсах із соціології культури, методології соціологічного дослідження, аналізу соціальних даних, контент-аналізу та цифрових гуманітарних студій. Окреме практичне значення має фіксація обмежень автоматизованого словникового підходу для подальшого вдосконалення україномовних ресурсів аксіологічного аналізу.

Апробація результатів магістерської дисертації. Окремі положення магістерської дисертації апробовано у підготовленій науковій статті «Методи аналізу ціннісних систем у художньому дискурсі: текстоцентричний підхід» (Маїлян, 2026), а також у тезах доповіді «Analysis of Value Narratives in Fiction as a Factor of Social Influence in the Digital Media Environment», представлених на III Міжнародній науково-практичній онлайн-конференції «Фундаментальні та прикладні проблеми суспільства: історія, сьогодення, майбутнє» (Київ, 16 квітня 2026 р.) (Mailian & Korzhov, 2026).

Публікації:

Стаття: Маїлян, А. С. (2026). Методи аналізу ціннісних систем у художньому дискурсі: текстоцентричний підхід. *Вісник науки та освіти. Серія: Філологія*. [https://doi.org/10.52058/2786-6165-2026-4\(46\)-1444-1491](https://doi.org/10.52058/2786-6165-2026-4(46)-1444-1491).

Тези конференції: Mailian, A., & Korzhov, H. (2026, 16 квітня). Analysis of value narratives in fiction as a factor of social influence in the digital media environment. У *Фундаментальні та прикладні проблеми суспільства: історія,*

сьогодення, майбутнє: Матеріали III Міжнародної науково-практичної онлайн-конференції (С. 57–60). Державний торговельно-економічний університет. <https://doi.org/10.31617/k.knute.2026-04-16>.

Тези конференції: Маїлян, А. С. (2025, 27–28 лютого). Комплексне застосування кількісних і якісних методів аналізу тексту в ЗМІ для визначення можливості впливу на настрій, емоційний стан та політичну спрямованість читачів. У *Україна у просторі соціальних комунікацій: суспільство, медіа, рефлексії війни: Матеріали конференції* (С. 672–675). Державний торговельно-економічний університет. <https://doi.org/10.31617/k.knute.2025-02-27-28>.

Структура магістерської дисертації. Магістерська дисертація складається зі вступу, чотирьох розділів, висновків, рекомендацій, списку використаних джерел і додатків. Загальний обсяг основного тексту роботи становить **118** сторінок. Список використаних джерел нараховує **158** найменувань. Робота містить **16** таблиць, **11** рисунків та **16** додатків (А–Т).

У першому розділі розглянуто теоретико-методологічні засади дослідження: понятійний апарат, соціологічні концепції цінностей, контекст соціальної трансформації українського суспільства 1991–2025 рр., специфіку художнього тексту як соціологічного джерела, текстоцентричний підхід до аналізу ціннісних репрезентацій, а також сформульовано теоретично мотивовані очікування (гіпотези Н1–Н4) щодо динаміки ціннісних репрезентацій, які перевіряються в емпіричній частині дослідження.

У другому розділі описано формування корпусу україномовної художньої прози 1991–2025 рр., жанрову класифікацію, часову періодизацію, побудову аналітичного датасету та технологічну реалізацію автоматизованого аксіологічного кодування.

У третьому розділі здійснено валідацію автоматизованого аксіологічного кодування — перевірку внутрішньої надійності, зовнішньої валідності та методичних меж використання отриманих результатів.

У четвертому розділі проаналізовано первинну динаміку ціннісних категорій у корпусних даних, здійснено жанрову стандартизацію, перевірено теоретичні гіпотези, розглянуто показові й контрастні кейси, виконано діагностичне зіставлення з даними WVS, EVS і національних опитувань; у висновках до розділу сформульовано узагальнену інтерпретацію трансформації ціннісних репрезентацій, у центрі якої — робочий патерн розходження між лексично-маркованим і декларативно-самозвітним режимами фіксації цінностей.

РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ ТРАНСФОРМАЦІЇ ЦІННІСНИХ ОРІЄНТАЦІЙ У ДИСКУРСІ УКРАЇНСЬКОЇ ХУДОЖНЬОЇ ЛІТЕРАТУРИ

1.1. Цінності, ціннісні орієнтації та ціннісні системи: понятійний апарат, теоретичні моделі і соціологічна концептуалізація

Поняття «цінність» у міждисциплінарному полі: термінологічна різнорідність і соціологічний ракурс.

Поняття «цінність» належить з одного боку до базових для суспільства, але з іншого боку, водночас і до найбільш неоднозначних категорій гуманітарного й соціального знання. У філософії, психології, соціології, культурології та педагогії це поняття використовується для позначення різних явищ: об'єктивної значущості, суб'єктивних переваг, нормативних стандартів, цілей та мотивацій або соціокультурних орієнтирів (Hitlin & Piliavin, 2004; Kluckhohn, 1951; Rohan, 2000). Для дослідження художнього дискурсу така багатозначність поняття має значення не лише термінологічне, а й методологічне - від прийнятого розуміння цінності буде залежати як вибір одиниць аналізу, так і логіка кодування та межі інтерпретації текстових маркерів.

Історично поняття цінності оформлювалося у кількох теоретичних лініях. У філософській аксіології другої половини XIX - початку XX ст. воно розглядалося переважно в опозиції між об'єктивістськими й суб'єктивістськими підходами. Неокантіанська традиція пов'язувала цінності з надіндивідуальними

значущостями, що не зводяться до психологічних станів суб'єкта (Rickert, 1921; Windelband, 1903). Феноменологічна аксіологія Scheler (1973), спираючись на проблематику емоційного осягнення, трактувала цінності як об'єктивні якості, що відкриваються у переживанні, тоді як натуралістична лінія пов'язувала цінність з інтересом суб'єкта. Для цієї роботи ця дискусія важлива не як самостійний предмет аналізу, а як підстава для розуміння того, чому цінності можуть інтерпретуватися або як незалежні нормативні орієнтири, або як результат людських оцінок і практик.

У соціально-психологічній та крос-культурній традиції акцент зміщується від онтологічного статусу цінностей до їхньої операціоналізації. Рокіч визначав цінності як стійкі переконання щодо бажаних способів поведінки або кінцевих станів існування (Rokeach, 1973). Він розмежовував інструментальні й термінальні цінності. Шварц, розвиваючи цю лінію, розглядав цінності як надситуативні цілі мотивації, які виконують функцію керуючих принципів у житті людини або групи і створюють структуровану систему сумісностей і конфліктів (Schwartz, 1992). Такі підходи до розуміння цінностей важливі для подальшої їх операціоналізації, однак самі по собі вони не вичерпують завдань цієї роботи. Підходи Рокіча та Шварца зосереджуються переважно на індивідуальному або крос-культурно-порівняльному рівні.

Соціологічний ракурс зміщує фокус аналізу. В межах такого фокусу цінність розглядається не як індивідуальне переконання чи переживання, а як елемент колективних нормативних систем, культурних класифікацій, соціальних практик та інституційного порядку (Hitlin & Piliavin, 2004). Для даної магістерської роботи саме така перспектива виступає теоретичним ядром. Її вибір зумовлений характером самого матеріалу та предмета дослідження. Корпус україномовної художньої прози 1991-2025 рр. - це не індивідуальні звіти опитуваних і не біографічні документи, а тексти, що функціонували у публічному культурному просторі, проходили процедури редактури, видання;

вони репрезентують скоріше колективно поділювані уявлення, ніж особисті переваги окремих суб'єктів. Трансформація ціннісних орієнтацій як предмет аналізу за визначенням передбачає осмислення на колективному та суспільному рівні, бо стосується змін у соціально розподілених, а не суто індивідуальних структурах розуміння важливого та не важливого. Зрештою, лише за умови розуміння цінностей як соціально й культурно опосередкованих орієнтирів, здатних репрезентуватися через символічні системи, художній дискурс взагалі може бути використаний як соціологічне джерело (Durkheim, 1912/1995).

Розуміння й прийняття соціологічного ракурсу як теоретичного ядра не передбачає відмови від суміжних дисциплін. Філософська аксіологія продовжує бути корисною для прояснення онтологічного статусу цінностей, а соціально-психологічні моделі – корисні для операціоналізації окремих ціннісних категорій. Однак рамку інтерпретації задає саме соціологія цінностей. У наступному підрозділі вона розгортається у двох блоках: класична традиція (Durkheim, 1912/1995; Weber, 1946a; Thomas & Znaniecki, 1918; Parsons, 1951), яка окреслила інституційно-нормативний рівень аналізу, і сучасні емпіричні моделі (Inglehart, 1977; Schwartz, 1992; Bourdieu, 1990), що дозволяють описувати ціннісні трансформації як структуровані, динамічні процеси. Саме з цієї позиції в наступному підрозділі розглядаються класичні й сучасні соціологічні концепції цінностей.

Ключові соціологічні концепції цінностей: від класичних підходів до сучасних емпіричних моделей.

Соціологічна традиція розглядає цінності не лише як особисті переконання або пріоритети мотивації, а як елементи культурного порядку, колективних уявлень, соціальних практик та норм інституцій (Hitlin & Piliavin, 2004). Для завдань цієї роботи важливим є не охоплення всіх можливих теорій цінностей, а

розуміння тих підходів, які дозволяють перейти від загального поняття «цінності» до аналізу саме текстової репрезентації її в корпусі україномовної художньої прози 1991-2025 рр. В дослідженні розгляд концепцій здійснюється за їхньою функцією - одні концепції задають розуміння цінностей як надіндивідуальних соціокультурних орієнтирів, інші - виконують допоміжну функцію при описі їхньої структури, динаміки або зв'язків із культурним полем.

Класична соціологічна традиція заклала розуміння цінностей як соціально вкорінених орієнтирів, що не зводяться до приватних уподобань. Аналізуючи колективну свідомість і колективні уявлення, Еміль Дюркгейм показав, що моральні смисли мають надіндивідуальний характер і забезпечують символічне розмежування значущого, належного, забороненого й буденного (Durkheim, 1893/1984, 1895/1982, 1912/1995). Така перспектива важлива в цьому дослідженні тому, що дає змогу розглядати літературні образи, конфлікти й моральні опозиції не лише як індивідуальні естетичні рішення автора, а як можливі форми репрезентації колективних уявлень.

Макс Вебер доповнює цю рамку через поняття ціннісно-раціональної дії та ціннісних сфер. Ціннісно-раціональна дія орієнтується на переконаність у значущості певного способу поведінки незалежно від безпосередньої вигоди (Weber, 1978). Концепція ціннісних сфер показує, що релігійна, економічна, правова, політична й естетична логіки можуть вступати між собою в напругу (Weber, 1946b). У контексті цього дослідження це дозволяє трактувати художній дискурс як простір, у якому суспільні ціннісні конфлікти отримують символічну й наративну форму.

Важливий для цієї роботи методологічний місток запропонували Вільям Томас і Флоріан Знанецький. У «Польському селянині в Європі та Америці» вони розмежували соціальну цінність як культурно значущий об'єкт або зміст і установку як індивідуальну орієнтацію щодо такого об'єкта (Thomas & Znaniecki, 1918). Їхнє використання особистих документів як матеріалу для

реконструкції соціальних змін є важливим прецедентом для текстоцентричного дослідження. Водночас художній твір не можна прирівнювати до прямого документа досвіду: він є естетично й жанрово опосередкованою формою репрезентації соціальних цінностей.

Толкотт Парсонс систематизував соціологічне розуміння цінностей у межах теорії соціальної дії та соціальних систем. Цінності у нього виступають елементами спільної символічної системи, що задають стандарти вибору між альтернативними орієнтаціями дії (Parsons & Shils, 1951). Їхня роль у підтриманні культурного зразка означає, що зміна цінностей зачіпає не окремі переконання, а ширшу конфігурацію соціального порядку (Parsons, 1951). Для цієї роботи це важливо як підстава аналізувати не тільки частоту окремих ціннісних понять у тексті, а й способи організації нормативних моделей належного.

Сучасні емпіричні моделі цінностей важливі передусім тому, що пропонують інструменти їхньої операціоналізації. Мілтон Рокіч розмежував термінальні цінності, тобто бажані кінцеві стани, та інструментальні цінності, тобто бажані способи поведінки (Rokeach, 1973). У роботі з художніми текстами це розмежування дає змогу відокремлювати цілі, які персонажі або наратор декларують як значущі, від способів дії, через які ці цілі реалізуються або проблематизуються.

Зв'язок змін цінностей із зміною в поколіннях пов'язаний із підходом Карла Маннгейма та теорією Рональда Інглхарта. Маннгейм показав, що покоління формується не лише біологічним віком, а спільним історичним досвідом (Mannheim, 1952). У самостійному теоретичному русі досліджень Інглхарт і Велзель розвинули цю логіку в моделі міжпоколінних ціннісних змін та через протиставлення цінностей виживання і самовираження (Inglehart, 1977; Inglehart & Welzel, 2005). Для дослідження української художньої прози ці моменти мають значення з боку розуміння зовнішнього соціологічного

контексту - дані досліджень WVS та EVS для України можуть використовуватися як порівняльний фон для аналізу ціннісних репрезентацій відповідних хронологічних зрізів (European Values Study, n.d.; Український центр європейської політики & Info Sapiens, 2020; World Values Survey Association, n.d.).

Найважливішою операціональною моделлю для цієї роботи є теорія базових людських цінностей Шалом Шварца. Його десять базових типів цінностей утворюють мотиваційне коло, у якому суміжні цінності є відносно сумісними, а протилежні - потенційно конфліктними (Schwartz, 1992). Подальше розширення моделі до 19 диференційованих цінностей уточнює можливості більш чутливого аналізу ціннісної структури (Schwartz et al., 2012). У цьому дослідженні модель Шварца розглядається не як готова процедура психометрії, а як теоретичне підґрунтя для опису змін ціннісних кластерів у художньому дискурсі.

Питання того, чому в художньому тексті конкретні ціннісні позиції взагалі стають помітними, а не залишаються маргінальними, лежить за межами вимірювальних моделей. Відповідь на питання пропонує П. Бурдьє через введення поняття хабітусу - системи стійких диспозицій, через які цінності відтворюються на практиці, в відтінках, смаках і судженнях без прямого усвідомленого проголошення (Bourdieu, 1990). З погляду теорії поля культурного виробництва художній текст формується не лише прямою авторською волею, а й позицією автора в літературному полі, жанровими очікуваннями, особливостями видавничих механізмів та структурами символічного поля (Bourdieu, 1993). Важливо розуміти, що ціннісний дискурс літератури не можна напряду ототожнювати з персональними поглядами автора або з безпосереднім «віддзеркаленням» суспільства.

Всі зазначені концепції сукупно утворюють багаторівневу рамку, у якій інституційно-нормативний вимір цінностей класичної традиції поєднується з

можливостями їх операціоналізації (Rokeach, 1973; Schwartz, 1992), динамікою зміни поколінь (Inglehart, 1977; Inglehart & Welzel, 2005; Mannheim, 1936) і загальною чутливістю до структури культурного поля (Bourdieu, 1990, 1993). У застосуванні до аналізу художнього дискурсу це означає, що окрема цінність у тексті ніколи не існує ізольовано: вона входить у ієрархію норм конкретного періоду, проявляється у конфлікті ціннісних сфер, операціоналізується через типи (термінальні/інструментальні, мотиваційне коло), має історичну поколінну обумовленість і пов'язана з існуванням літератури в культурному полі. Подальші підрозділи конкретизують цю рамку через розмежування ключових понять і способів їх виявлення в художньому тексті. Деталізоване зіставлення концепцій подано в Додатку А.

Розмежування понять: «цінності», «ціннісні орієнтації», «ціннісна система», «ціннісні установки»

Дослідження трансформації ціннісних орієнтацій у художньому дискурсі потребує не лише загального визначення поняття «цінність», а й розмежування кількох близьких категорій. У цій роботі ключовими є чотири поняття: «цінність», «ціннісна орієнтація», «ціннісна система» та «ціннісна установка». Вони описують різні рівні ціннісної регуляції: від загального соціокультурного орієнтира до його привласнення суб'єктом, структурної організації й ситуативної актуалізації. Їхнє змішування створює ризик підміни аналізу стійких ціннісних конфігурацій аналізом окремих реакцій, норм або переконань.

Цінність у межах цього дослідження розуміється як соціально й культурно опосередкований надситуативний орієнтир оцінки та дії, що задає критерії розмежування бажаного й небажаного, належного й неприйняттого, значущого й другорядного. Таке розуміння спирається на спільні ознаки, що фіксуються у впливових операціоналізованих підходах: відносну стійкість, надситуативність і

здатність виконувати функцію критерію оцінювання (Hitlin & Piliavin, 2004; Rokeach, 1973; Schwartz, 1992). У художньому тексті цінність рідко виявляється лише через пряме називання. Частіше вона реконструюється через оцінну лексику, моральні опозиції, повторювані образи, конфлікти, авторську або нараторську модальність і способи сюжетного винагородження чи засудження певних дій.

Ціннісна орієнтація позначає спосіб індивідуального або групового привласнення й ранжування цінностей. Якщо цінність є загальним соціокультурним орієнтиром, то ціннісна орієнтація показує, яким саме цінностям суб'єкт або група надає пріоритет у конкретному контексті. Клакхон і Стродтбек розглядали ціннісні орієнтації як ієрархічно впорядковані принципи, що поєднують когнітивний, афективний і директивний компоненти оцінювання та спрямовують дію (Kluckhohn & Strodtbeck, 1961). Для цієї роботи важливо, що йдеться не про психометричне вимірювання реальних індивідів, а про реконструкцію стійких пріоритетів, репрезентованих у позиціях персонажів, наратора або ширших текстових моделях належного.

Ціннісна система фіксує не окрему цінність і не тільки індивідуальний профіль її привласнення, а структуровану конфігурацію взаємопов'язаних цінностей. У Рокіча ціннісна система пов'язана з організацією вибору між конкуруючими цінностями (Rokeach, 1973), а у Шварца — з мотиваційним колом, де суміжні цінності є відносно сумісними, а протилежні — потенційно конфліктними (Schwartz, 1992). Тому зміна ціннісної системи означає не лише появу або зникнення окремих слів-маркерів, а переієрархізацію всієї конфігурації значущого. У художньому дискурсі така система реконструюється через повторювані моральні опозиції, сюжетні розв'язання конфліктів, готовність персонажів жертвувати одними благами заради інших і стійкі моделі оцінювання соціальної реальності.

Ціннісна установка позначає більш ситуативний рівень ціннісної регуляції - готовність суб'єкта до певної оцінки, реакції або дії щодо конкретного об'єкта чи ситуації. Вона спирається на ціннісні орієнтації суб'єкта, але не зводиться до них, оскільки залежить від контексту, інтерпретації ситуації та переконань. Класичне розмежування соціальної цінності й установки було запропоноване Томасом і Знанецьким: цінність постає як культурно значущий об'єкт, а установка — як індивідуальна орієнтація щодо нього (Thomas & Znaniecki, 1918). У соціально-психологічній традиції установка також описується через когнітивний, афективний і поведінковий компоненти (Fishbein & Ajzen, 1975; Rosenberg & Hovland, 1960). У художньому тексті ціннісні установки виявляються через ситуативні реакції персонажів, модальність висловлювань, форми виправдання або осуду, емоційно забарвлені оцінки й готовність діяти певним чином у конкретних обставинах.

Поряд із цими чотирма поняттями у дослідженні враховуються суміжні категорії, що описують або нормативно-інституційне середовище ціннісної регуляції, або її мотиваційно-когнітивні передумови. До першої групи належать **норми** — правила очікуваної поведінки, що конкретизують цінності в конкретному інституційному контексті (Parsons, 1951), — та **ідеологія**, яка організує групові переконання, цінності й пояснювальні схеми, легітимізуючи або оспорюючи соціальний порядок (van Dijk, 1998).

Другу групу утворюють поняття, що описують суб'єктивні передумови дії. **Потреби** позначають мотиваційні стани нестачі або розвитку, які спонукають до дії, але не зводяться до ціннісних орієнтацій (Maslow, 1954). **Інтереси** пов'язані зі ставками, ресурсами й позицією суб'єкта у соціальному або культурному полі (Bourdieu, 1990, 1993). **Переконання** ж фіксують когнітивний вимір того, що суб'єкт вважає істинним або правильним; цінності є особливим типом переконань, пов'язаним не лише з істинністю, а з бажаним і належним (Rokeach,

1973). Розгорнуте зіставлення усіх дев'яти понять разом із можливими текстовими індикаторами подано в Додатку Б.

У подальшому дослідженні цінності розуміються як соціально й культурно опосередковані орієнтири оцінки, вибору та дії, що можуть репрезентуватися у художньому дискурсі через лексичні маркери, моральні опозиції, сюжетні конфлікти, образи персонажів і моделі належного. Ціннісні орієнтації трактуються як відносно стійкі способи ієрархізації таких орієнтирів на індивідуальному, груповому або дискурсивному рівні, а ціннісна система — як структурована конфігурація взаємопов'язаних цінностей, що може змінюватися історично. Для емпіричної частини роботи це означає, що аналіз спрямований не на пряме вимірювання переконань авторів або читачів, а на виявлення динаміки текстових репрезентацій ціннісно маркованих смислів у корпусі україномовної прози.

1.2. Соціальна трансформація і ціннісні зрушення в постсоціалістичному суспільстві: теоретичний вимір та український контекст 1991–2025 рр.

Соціальна трансформація як теоретична рамка дослідження ціннісних змін.

Під «соціальною трансформацією» в цій роботі розуміється не будь-яка зміна в соціумі, а така перебудова суспільної системи, яка одночасно зачіпає і соціальні інститути, і соціальні практики, і культурні норми, ідентичності та способи оцінювання бажаного й належного. Якщо реформа зазвичай має частковий і керований характер, то трансформація є багаторівневим і не завжди односпрямованим процесом. Також вона не зводиться до модернізації, оскільки

не обов'язково означає лінійне наближення до наперед заданої моделі розвитку (Sztompka, 1993a). Саме тому поняття трансформації є продуктивним для аналізу України 1991-2025 рр., адже в цей період інституційні, культурні й ціннісні зміни відбувалися нерівномірно, через кризи, повернення до попередніх моделей і через нові хвилі переозначення соціального порядку.

У логіці Штомпки соціальна трансформація може бути описана як зміна кількох взаємопов'язаних рівнів суспільного життя: структурного, інституційного, культурного та суб'єктивно-свідомісного (Sztompka, 1993a). Для цього дослідження принципово важливо, що ціннісний рівень не є лише наслідком інституційних змін. Неформальні норми, моральні критерії, ієрархії символів та повсякденні моделі оцінювання змінюються в суспільстві повільно або суперечливо і неоднозначно. На відміну від них, формальні правила можуть бути змінені швидко. Така різниця між темпами інституційних змін і змін у неформальних ціннісних нормах в цій роботі позначається як «аксіологічний лаг». Це робоче поняття спирається на ідею Огберна про культурний лаг, конкретизуючи її щодо сфери цінностей (Ogburn, 1922). Аксіологічний лаг у художньому дискурсі може проявлятися як одночасне існування «застарілих» моральних опозицій із новими ціннісними вимогами, існування нерозв'язаних суперечностей між нормами попередньої епохи й новим етичним словником.

У постсоціалістичному контексті існування країни аксіологічний лаг набуває особливої ваги. Руйнування попереднього суспільного порядку не означає автоматичного виникнення нової узгодженості цінностей. Українські соціологи Головаха та Паніна описували подібний стан через проблематику нормативної дезорієнтації, деморалізації та деінституціоналізації. В такі моменти старі правила вже втрачають регулятивну силу, а нові ще не набувають стабільності й довіри суспільства (Головаха & Паніна, 1994, 2001). Схожу логіку має поняття культурної травми Штомпки - радикальна й раптова зміна в суспільстві може переживатися як розрив культурної тканини цього суспільства,

що вже зачіпає ідентичності, довіру, пам'ять і способи колективного осмислення майбутнього (Sztompka, 2000). Для аналізу художнього дискурсу це означає, що література може фіксувати не лише поступову зміну ціннісних пріоритетів, а й символічне переживання втрати нормативної опори.

Розгляд змін у системі цінностей потребує також урахування логіки часу і зміни поколінь. У постматеріалістичній теорії Інглхарт пов'язував зміну цінностей із гіпотезами нестачі та соціалізації: умови безпеки або дефіциту, пережиті в ранньому життєвому досвіді, впливають на подальші ціннісні пріоритети поколінь (Inglehart, 1997). При цьому українська траєкторія змін суспільства не може бути описана просто як поступовий перехід від матеріалістичних до постматеріалістичних. Економічні, політичні й безпекові кризи, повторюючись, здатні поновити актуальність цінностей виживання навіть тоді, коли водночас зростає значущість свободи, гідності, автономії або національної суб'єктності (European Values Study, n.d.; Український центр європейської політики & Info Sapiens, 2020; World Values Survey Association, n.d.).

Для подальшого аналізу у роботі пропонується розрізняти три типи ціннісних змін. **Ціннісна криза** означає стан дезорієнтації, за якого попередні орієнтири втратили переконливість, а нові ще не стали загальновизнаними. **Ціннісне зрушення** позначає поступову зміну в ієрархії вже наявних цінностей без повного руйнування системи. **Ціннісна реконструкція** передбачає активне формування нової або оновленої схеми відносин через задіяння ресурсів пам'яті, ідентичності, досвіду криз або під зовнішнім тиском, зокрема воєнним. Логіка розмежування цих типів ціннісних змін спирається на загальне розуміння процесу соціологічного аналізу трансформацій, де етапи дестабілізації, перерозподілу значущого та формування нового порядку можуть розрізнятися як аналітично, так і в часовому вимірі (Sztompka, 1993a, 2000). Далі поняття «криза», «зрушення», «реконструкція» використовуються як аналітична сітка

для інтерпретації чотирьох хронологічних зрізів українського дискурсу художньої літератури 1991-2025 рр.

Постсоціалістичне суспільство як простір ціннісної переієрархізації: теоретичні закономірності та емпіричні індикатори.

Постсоціалістичне суспільство в цьому дослідженні розглядається не лише як суспільство після зміни політичного режиму, а як простір тривалої ціннісної переієрархізації. Йдеться не про одномоментне заміщення однієї системи норм іншою, а про ситуацію, у якій державно-соціалістичний нормативний порядок швидко втрачає легітимність, тоді як нові орієнтири — демократія, ринок, автономія, права людини, громадянська участь, національна суб'єктність — укорінюються нерівномірно. Цю складність добре пояснює концепція Оффе про «потрійний перехід»: після 1989-1991 рр. країни Центральної та Східної Європи одночасно мали розв'язувати завдання державного будівництва, демократизації та переходу до ринкової економіки (Offe, 1991). Унаслідок цього ціннісна система зазнавала тиску одразу з кількох напрямів - із політичного, економічного, з боку інституцій та з боку процесів ідентифікації.

Одну з ключових проблем посткомуністичних суспільств Штомпка описував через поняття «цивілізаційної некомпетентності» як дефіциту культурних правил, звичок і практичних диспозицій, необхідних для участі в модерних демократичних, ринкових і громадянських інститутах (Sztompka, 1993). У цій роботі це поняття використовується не як оцінка рівня «відсталості», а як аналітичне позначення розриву між формально встановленими інститутами та ціннісними схемами поведінки у повсякденні, що були сформованими попереднім досвідом. Нові інститути вимагали наявності певної довіри до незнайомих, готовності до самоорганізації та прийняття ризику й відповідальності. В той самий час радянська спадщина часто підтримувала

патерналістські очікування, недовіру до публічних ініціатив і спирання на приватні неформальні мережі (Pop-Eleches & Tucker, 2017; Sztompka, 1993).

Емпіричним проявом цього розриву є слабкість організованої громадянської участі, проаналізована Говардом на матеріалі посткомуністичної Європи. На його думку, низький рівень членства у добровільних організаціях пов'язаний із досвідом формально контрольованої участі за соціалістичного ладу, збереженням щільних приватних мереж і розчаруванням у результатах трансформації (Howard, 2002, 2003). Для аналізу цінностей це означає, що постсоціалістична переієрархізація проявляється не тільки в деклараціях, а й у практиках: солідарність часто залишається приватизованою, довіра — обмеженою, а публічна участь — нестійкою.

Компаративні дані WVS та EVS дають змогу окреслити загальний фон цих змін. У 1990-х - на початку 2000-х рр. багато постсоціалістичних суспільств тяжіли до конфігурації «секулярно-раціональні цінності / цінності виживання»: спадщина комуністичної модернізації поєднувалася з матеріальною нестачею, низькою довірою та слабкою індивідуальною автономією (European Values Study, n.d.; Inglehart & Welzel, 2005; World Values Survey Association, n.d.). Водночас рух до цінностей самовираження був повільним і нерівномірним, що не узгоджувалося з лінійними моделями модернізації (Inglehart & Welzel, 2005; European Values Study, n.d; World Values Survey Association, n.d.). Тому WVS та EVS у цій роботі використовуються як зовнішня порівняльна рамка, а не як прямий інструмент вимірювання цінностей у художньому тексті.

На рівні емпіричних індикаторів постсоціалістична переієрархізація може фіксуватися через міжособистісну й інституційну довіру, готовність до участі в добровільних організаціях, ставлення до автономії та відповідальності, толерантність до інакшості, оцінку ролі держави, значущість національної ідентичності й поколінні відмінності (Howard, 2002; Inglehart & Welzel, 2005; Mishler & Rose, 2001; World Values Survey Association, n.d.). Однак у художньому

дискурсі ці індикатори не з'являються у формі опитувальних змінних. Вони можуть бути реконструйовані через мотиви недовіри, приватизації життєвого світу, патерналістських очікувань, моральної амбівалентності, конфлікту безпеки й свободи, переозначення мови, пам'яті, держави, Заходу та громадянської дії.

Напрямки руху у процесі постсоціалістичних змін суттєво відрізнялися між країнами. Євроінтеграція, швидша демократизація та наявність дієздатного громадянського суспільства створювали сприятливіші умови для плюралізації публічного простору та посилення орієнтацій на автономію. Це процеси виразніше простежувалися в Польщі, країнах Балтії та Чехії (Berglund et al., 2013; Vachudova, 2005). Інакше формувалася динаміка в суспільствах із глибшою інституційною невизначеністю й конкуруючими геополітичними орієнтаціями. Там була помітна нелінійна динаміка у поєднанні зрушень лібералізації, патерналістських реверсій, низької довіри у суспільстві та кризових сплесків національної мобілізації (Brubaker, 1996). Саме до такого типу належить український випадок, який потребує окремого аналізу в наступному підрозділі.

Ціннісні трансформації в Україні (1991-2025): рушійні сили, кризові події та хронологічні зрізи дослідження.

Траєкторія руху України у постсоціалістичних трансформаціях відрізняється від частини країн Центрально-Східної Європи більшою нелінійністю, тривалою інституційною невизначеністю та залежністю від кризових подій. Цю траєкторію не доцільно описувати просто як поступову заміну «старої» системи цінностей «новою». Скоріше у випадку України йдеться про послідовність переозначень, під час яких соціально-економічні кризи, громадянські мобілізації, конкуренція геополітичних орієнтацій і російська агресія змінювали саму логіку ціннісного вибору. У цьому контексті важливими

рушіями змін виступали пострадянська деінституціоналізація і нормативна невизначеність (Головаха & Паніна, 2001), низька інституційна довіра (Mishler & Rose, 2001; Головаха et al., 2014), поступове формування громадянського суспільства (Howard, 2002; 2003), конкуренція європейського, пострадянського й національно-демократичного політичних проєктів (Kuzio, 2010; Vachudova, 2005). Також значну роль зіграли воєнні події 2014 і 2022 рр., які радикально перевизначили співвідношення цінностей безпеки, свободи, гідності, національної ідентичності та солідарності (Бакіров & Сокур'янська, 2023; Головаха & Макєєв, 2022; Grushetsky & Paniotto, 2025).

Для подальшого проведення корпусного аналізу ця траєкторія поділяється на чотири хронологічні зрізи. Межування цих зрізів виконується не лише календарно, а й через фіксацію критичних подій, що змінювали інтенсивність і спрямованість динаміки ціннісних змін у суспільстві.

Перший зріз – 1991-2003 рр. — характеризується пострадянською аномією і переважанням цінностей виживання. Розпад радянської нормативної системи, економічна дезорганізація, гіперінфляція, зубожіння й недовіра до політичних інститутів створили такі умови, у яких особисте виживання часто ставало важливішим за публічну участь (Головаха & Паніна, 1994, 2001; Головаха, 2016). Дані досліджень WVS/EVS середини 1990-х рр. і їх подальші інтерпретації в межах моделі Інглхарта-Велзеля узгоджуються з характеристикою України як суспільства з вираженою орієнтацією на виживання (Inglehart & Welzel, 2005; World Values Survey Association, n.d.-b). Водночас уже в цей період виникали нерівномірні «острівці» нових орієнтацій - підприємницьких, індивідуалістичних, прозахідних, молодіжних, - які не скасовували загальної аномічності, але формували передумови майбутньої поляризації.

Другий зріз - 2004-2013 рр. - пов'язаний із ціннісною поляризацією та хвилями громадянської мобілізації. Помаранчева революція актуалізувала цінності чесних виборів, справедливості, гідності й демократичного

представництва, але ці очікування не отримали стійкого інституційного закріплення. Дані WVS 2006 р. дають підстави говорити про певне посилення орієнтацій самовираження, тоді як WVS 2011 р. фіксує збереження регіональної й поколінної неоднорідності (World Values Survey Association, n.d.-c, n.d.-d; Український центр європейської політики & Info Sapiens, 2020). Для цього періоду характерне співіснування громадянського піднесення, розчарування результатами революції та конкуренції проєвропейського й пострадянського векторів.

Третій зріз – 2014-2021 рр. - визначається Революцією Гідності, початком російської агресії та прискореним переозначенням ідентичності. Анексія Криму й війна на Донбасі посилили значущість концепцій гідності, безпеки, національної суб'єктності, посилилася роль української мови та важливість дистанціювання від російського культурно-політичного простору (Kulyk, 2016). Дані WVS 2020 р. засвідчують рух у напрямі секулярно-раціональних цінностей і самовираження, але їх слід інтерпретувати обережно - вибірка не охоплювала Автономну Республіку Крим і частину Донбасу. В цих умовах порівняння з попередніми хвилями частково відображає зміну складу генеральної сукупності (Український центр європейської політики & Info Sapiens, 2020). Для художнього дискурсу цього періоду особливо важливими стають маркери пам'яті, мови, війни, гідності й межі між «своїм» та «чужим».

Четвертий зріз – це період від початку 2022 по 2025 рр. включно. Цей зріз відзначений процесами екзистенційної мобілізації після повномасштабного вторгнення. Український випадок ускладнює класичне припущення постматеріалістичної теорії: фізична небезпека не лише реактуалізує цінності виживання, а й поєднується з високою значущістю свободи, державної незалежності, солідарності, громадянської участі та готовності до опору (Бакіров & Сокур'янська, 2023; Головаха & Макеев, 2022; Grushetsky & Paniotto, 2025). Водночас воєнна консолідація не усуває внутрішніх напружень: зберігається

запит на справедливість, оновлення еліт, подолання старих політичних практик і переосмислення відносин між особистим та колективним (Головаха & Макєєв, 2023; Київський міжнародний інститут соціології, 2024; Разумков Центр, 2026).

Запропонований поділ на чотири зрізи виконує подвійну функцію - зрізи фіксують історичну періодизацію постнезалежного періоду і водночас створюють аналітичну рамку для подальшого дослідження художнього дискурсу. У термінах підрозділу 1.2 перший період відповідає переважно ціннісній кризі; другий - поєднанню зрушення й поляризації; третій - прискореному перевизначенню ідентичностей; четвертий - активній ціннісній реконструкції в умовах екзистенційної загрози. На рівні корпусу української прози цій періодизації відповідає очікувана зміна домінуючих текстових маркерів: від мотивів виживання, недовіри 1990-х, через конфлікти громадянської участі й розчарування 2000-х, перевизначення мови, пам'яті й сприйняття війни в період після Євромайдану, до формування ціннісного блоку безпеки, свободи, солідарності й вимоги справедливості після 2022 р.

Соціальний контекст як інтерпретаційна рамка аналізу художнього дискурсу.

Теоретичні положення підрозділу 1.2 задають не лише історико-соціологічний фон дослідження, а й інтерпретаційну рамку подальшого аналізу художнього дискурсу. Соціальний контекст у цій роботі розуміється як сукупність криз, інституційних змін, ціннісних переієрархізацій і колективних уявлень, у межах яких художні тексти створюються, публікуються, читаються й набувають соціального значення. Соціальний контекст не варто розглядати як зовнішній чинник або причину, що механічно формує зміст літератури. Художній текст має власну естетичну логіку, свої жанрові обмеження й відносну автономію формування смислів.

За такого підходу важливо розрізнити й поєднати два теоретичні акценти. Концепція поля культурного виробництва Бурдьє дозволяє розглядати художній твір як продукт літературного поля, у якому діють авторські позиції, жанрові очікування, видавничі механізми, форми визнання та символічного капіталу (Bourdieu, 1993). Семіотична перспектива Лотмана наголошує, що художній текст не лише передає готові смисли, а моделює культурну реальність через образи, опозиції, норми й оцінки (Lotman, 1977, 1990). Використання цих концепцій у поєднанні дозволяє розглядати літературу одночасно як соціально зумовлений продукт і як простір символічного осмислення дійсності.

У межах цього дослідження соціальний контекст виконує дві взаємопов'язані функції: інтерпретаційну й порівняльну, що доповнюють одна одну. Як ресурс для інтерпретацій він задає умови коректного прочитання ціннісних маркерів. Одна й та сама лексема, сюжетна опозиція або модель поведінки набуває різного наповнення і смислу в ситуації пострадянської аномії 1990-х, після Помаранчевої революції, у період після 2014 р. або в умовах повномасштабної війни. Як ресурс для порівняння контекст забезпечує співвіднесеність хронологічних зрізів - зміни в художньому дискурсі набувають соціологічного значення лише тоді, коли вони співвіднесені з конкретними фазами трансформації суспільства. Поєднання зазначених функцій узгоджується з культурно-соціологічним підходом до літератури, запропонованим Ванею, який розглядає художню прозу як простір формування смислів, який виникає під час взаємодії тексту, читача й соціально-історичного середовища (Váňa, 2020).

Художній дискурс не слід зводити до пасивного відображення соціальної реальності. У перспективі критичних дискурс-досліджень дискурс може не лише репрезентувати, а й відтворювати або проблематизувати соціальні уявлення, ієрархії та владні відносини (van Dijk, 2008). Для цієї роботи це означає, що ціннісні маркери в художніх текстах розглядаються не як прямі показники

«реальних» цінностей населення, а як соціально й естетично опосередковані репрезентації, що потребують зіставлення з ширшим історичним контекстом.

Встановлена в підрозділі 1.2 чотириперіодна структура — 1991–2003, 2004–2013, 2014–2021, 2022–2025 рр. — у подальшому дослідженні виконує роль операціоналізованої моделі соціального контексту. Вона дозволяє співвідносити виявлені в корпусі ціннісні маркери з фазами пострадянської невизначеності, громадянської мобілізації, воєнно-ідентичнісного переозначення та екзистенційної ціннісної реконструкції. Така операціоналізація не заміщує власне аналізу художнього тексту, а створює зовнішню рамку, у межах якої можна оцінювати специфіку, інтенсивність і напрям виявлених у літературі ціннісних зрушень. Методологічні умови використання художнього тексту й художнього дискурсу як соціологічного джерела розглядаються в наступному підрозділі.

1.3. Художній текст і художній дискурс як соціологічне джерело дослідження цінностей

Художній текст і художній дискурс: розмежування понять у соціологічній перспективі.

Коректне розмежування понять «художній текст» і «художній дискурс» є необхідною умовою подальшого аналізу ціннісних орієнтацій у корпусі україномовної прози. У цій роботі ці поняття не вживаються як взаємозамінні: художній текст позначає конкретний твір як безпосередню одиницю корпусу й аналізу, тоді як художній дискурс охоплює ширший соціокультурний простір виробництва, циркуляції, рецепції та інтерпретації таких творів. Це розмежування визначає, що саме кодується в емпіричній частині дослідження і

на якому рівні отримані текстові маркери можуть бути соціологічно інтерпретовані.

У лінгвістиці текст зазвичай розглядається як смислова, структурна й комунікативна цілісність. Галлідей і Гасан визначали текст передусім як семантичну одиницю, тобто одиницю значення, а не лише форми (Halliday & Hasan, 1976). Де Богранд і Дресслер конкретизували текстуальність через стандарти зв'язності, когерентності, інтенціональності, прийнятності, інформативності, ситуативності та інтертекстуальності (de Beaugrande & Dressler, 1981). Для цього дослідження такі підходи важливі тому, що дозволяють розглядати художній твір як структурований матеріал для аналізу лексичних, наративних, композиційних та оцінних маркерів цінностей.

У соціологічній перспективі художній текст не зводиться до формально організованого мовного об'єкта. У семіотиці Ю. Лотмана текст постає як механізм культурного моделювання, здатний не лише передавати наявні смисли, а й організовувати та породжувати нові (Lotman, 1977, 1990). У бахтінській традиції романний текст розуміється як діалогічний і гетероглосний простір взаємодії різних соціальних голосів, мовних реєстрів і ціннісних позицій (Bakhtin, 1981). Тому аксіологічний аналіз художнього тексту не може обмежуватися пошуком прямої «авторської позиції»: ціннісний зміст твору формується через взаємодію наратора, персонажів, жанрових конвенцій, конфліктів і культурних кодів.

У межах цієї роботи художній текст розуміється як конкретний прозовий твір із визначеними текстуальними межами, внутрішньою структурою та зафіксованим місцем у корпусі. Саме він є одиницею відбору, кодування й первинного емпіричного аналізу. На рівні тексту фіксуються оцінна лексика, модальність, символічні образи, моральні опозиції, репліки персонажів, наративні конфлікти й інші маркери, через які цінності стають доступними для аналізу.

Поняття «дискурс» має ширший зміст. У соціогуманітарних дослідженнях воно позначає не лише мовлення або сукупність текстів, а й історично зумовлений порядок виробництва смислів, правил висловлювання та легітимації певних позицій (Titscher et al., 2000; Foucault, 1972). У критичному дискурс-аналізі, зокрема в моделі Феркло, дискурс розглядається як взаємозв'язок тексту, дискурсивної практики та соціальної практики (Fairclough, 1992). Ван Дейк додатково наголошує, що дискурсивні структури можуть брати участь у відтворенні ідеологій і владних відносин (van Dijk, 2008). Отже, дискурс виходить за межі окремого твору і включає умови його виробництва, поширення, прочитання та соціальної значущості.

У цьому дослідженні художній дискурс конкретизується як соціокультурний простір україномовної художньої прози 1991–2025 рр., що функціонував у межах чотирьох фаз соціальної трансформації, окреслених у підрозділі 1.2. На цьому рівні аналізу окремі текстові маркери співвідносяться з ширшими контекстами: літературним полем, жанровими очікуваннями, постколоніальними й транзитними процесами, змінами публічного мовлення про пам'ять, мову, війну, гідність, свободу та справедливість (Bourdieu, 1993; Гундорова, 2012). Відповідно, художній текст є рівнем фіксації ціннісних маркерів, а художній дискурс — рівнем їхньої соціологічної контекстуалізації.

Таке розмежування дозволяє уникнути двох методологічних помилок: редукції дослідження до імманентного аналізу твору без урахування соціального контексту і, навпаки, розчинення конкретного тексту в надто широкому понятті дискурсу без чітких одиниць аналізу. У подальшій логіці роботи текст забезпечує емпіричну фіксацію ціннісних маркерів, тоді як дискурс дозволяє інтерпретувати їх як соціально й культурно опосередковані репрезентації. Умови, можливості й обмеження такого використання художнього дискурсу розглядаються в наступному підрозділі.

Художній дискурс як джерело соціологічного вивчення ціннісних репрезентацій: обґрунтування, можливості та обмеження.

Виконане в попередньому підрозділі розмежування художнього тексту і художнього дискурсу ставить ключове методологічне питання - чи може художній дискурс бути коректним джерелом соціологічного дослідження ціннісних репрезентацій? Художня література не реалізує пряме відображення суспільної свідомості. Вона не може розглядатися як статистично репрезентативний зріз ціннісних орієнтацій населення. При цьому вона все ж таки фіксує в собі соціально значущі уявлення, моральні конфлікти, моделі належного й неналежного, символічні опозиції та способи культурного осмислення історичного досвіду. З цих причин художній дискурс в цій роботі розглядається як специфічне, самостійне, але не самодостатнє джерело соціологічного аналізу репрезентацій цінностей (Becker, 2007; Coser, 1963).

Теоретичне обґрунтування такого підходу спирається на культурну соціологію, соціологію літератури та дискурс-аналіз. У «сильній програмі» культурної соціології Александера культура не зводиться до похідної від соціальної структури, а розглядається як відносно автономна система смислів, здатна організовувати колективний досвід (Alexander, 2003). У цій логіці художній текст не є лише ілюстрацією соціальних процесів: він може символічно оформлювати культурні коди, моральні опозиції та фонові схеми колективного смислотворення. Доповнюючи цю позицію, Лонго показує спорідненість соціології та художньої прози через їхній спільний інтерес до соціальних відносин, мотивації дії, нормативних обмежень і повсякденних форм взаємодії (Longo, 2016). Концепція «культурного ромба» Грізволд уточнює цю позицію, розміщуючи художній текст між соціальним світом, творцем, рецептором і самим культурним об'єктом (Griswold, 2013). У цій моделі літературний твір є соціологічно інформативним саме тому, що перебуває в зоні перетину

авторських інтенцій, читацьких очікувань, інституцій літературного поля та ширших соціокультурних умов.

Особливу цінність художнього дискурсу для цього дослідження становить доступ до не лише декларованих, а й латентних форм ціннісної організації досвіду. Й. Ваня наголошує, що культурно-соціологічне дослідження літератури дозволяє виявляти ті виміри соціального знання, які втрачаються при прямому перекладі художнього смислу в соціологічні категорії (Váňa, 2020). Для аналізу цінностей це принципово: ціннісні орієнтири в художньому тексті часто не проголошуються прямо, а виявляються через конфлікти, пресупозиції, модальність, образи, сюжетні наслідки та мовчазні уявлення про нормальне, бажане або неприйнятне. Саме тому художній дискурс може доповнювати опитувальні дані, не дублюючи їх і не претендуючи на пряме вимірювання суспільних цінностей.

Методологічно таке використання потребує багаторівневої логіки інтерпретації. Спираючись на тривимірну модель Феркло, у роботі розрізняються текстовий, дискурсивний і соціальний рівні аналізу (Fairclough, 1992). На текстовому рівні фіксуються мовні, сюжетні та наративні маркери цінностей; на дискурсивному - вони співвідносяться з жанровими конвенціями, культурними кодами й позиціями літературного поля; на соціальному - обережно інтерпретуються як індикатори ширших ціннісних конфігурацій певного історичного періоду. Хоча така послідовність не усуває ризику помилкової інтерпретації - вона знижує небезпеку прямого ототожнення художнього тексту з емпіричним «зрізом» суспільної свідомості.

Можливості художнього дискурсу бути джерелом соціологічного дослідження цінностей полягають передусім у використанні діяхронного співставлення текстів різних періодів. Важливим також є виявлення повторюваних моральних опозицій і реконструкції способів символічного оформлення свободи, безпеки, гідності, справедливості, спільноти, пам'яті або

національної ідентичності (Longo, 2016; Griswold, 2013). Водночас ці можливості мають чіткі межі. Художній текст являє собою естетично опосередкований конструкт, який стилізує, ущільнює та іноді алегоризує соціальну реальність (Longo, 2016). Він формується в літературному полі з його жанровими конвенціями, видавничими механізмами, ринковими умовами, авторськими позиціями та структурами символічного капіталу (Bourdieu, 1993). Крім того, корпус художніх текстів не є статистично репрезентативним для всіх соціальних груп, а між часом написання, публікації й рецензії твору може існувати часовий лаг (Griswold, 2013; Longo, 2016).

Сукупність зазначених можливостей і обмежень означає, що для соціологічного використання художнього дискурсу потрібно використовувати не пряме перенесення висновків із тексту на суспільство, а будувати проміжну аналітичну ланку. Такою ланкою в цьому дослідженні є поняття ціннісної репрезентації - воно позначає спосіб, у який певні цінності символічно оформлюються, повторюються, конфліктують, легітимізуються або проблематизуються в корпусі художньої прози (Alexander, 2003; Longo, 2016; Váňa, 2020; Fairclough, 1992). Саме аналіз стійких репрезентацій у їх співвіднесенні із хронологічними зрізами 1991-2003, 2004-2013, 2014-2021 та 2022-2025 рр., дозволяє використовувати україномовну художню прозу як соціологічно інформативний матеріал для дослідження трансформації ціннісних орієнтацій. Конкретні рівні та форми таких репрезентацій розглядаються в наступному підрозділі.

Рівні репрезентації цінностей у художньому тексті: лексико-семантичний, наративний, дискурсивний та інтертекстуальний виміри.

Після обґрунтування художнього дискурсу як джерела соціологічного вивчення ціннісних репрезентацій необхідно уточнити, на яких рівнях

художнього тексту ці репрезентації можуть бути виявлені. У цьому дослідженні виокремлюються чотири взаємопов'язані виміри: лексико-семантичний, наративний, дискурсивний та інтертекстуальний. Таке розмежування спирається відповідно на аналіз оцінності, наратологію, критичний дискурс-аналіз і теорію інтертекстуальності (Fairclough, 1992; Genette, 1997; Hunston & Thompson, 2000; Kristeva, 1980; Rimmon-Kenan, 2002). Воно не означає поділу тексту на ізольовані шари: одна й та сама цінність може одночасно проявлятися через слово, сюжетний вибір, позиціювання соціальних акторів і апеляцію до культурної пам'яті.

Лексико-семантичний рівень є найбільш доступним для фіксації та систематичної обробки. На цьому рівні цінності виявляються через оцінну лексику, модальні конструкції, інтенсифікатори, заперечення та стійкі ціннісні концепти, такі як «свобода», «гідність», «батьківщина», «зрада». Оцінність у цьому контексті доцільно розуміти як мовне вираження позиції щодо описуваного явища (Hunston & Thompson, 2000). Водночас частотність слова не є достатнім доказом домінування відповідної цінності: важливо враховувати сполучуваність, модальність, позицію мовця, іронічний або полемічний контекст і те, чи подається певний концепт як схвалюваний, проблематизований або дискредитований.

Наративний рівень дозволяє реконструювати цінності через організацію подій, конфліктів, фокалізації та системи персонажів. Наратив завжди містить оцінний компонент, оскільки пояснює, чому певна подія є значущою і як її слід інтерпретувати (Labov & Waletzky, 1967). У художній прозі цей рівень доцільно аналізувати з урахуванням структурної наратології: ціннісний зміст проявляється через моральний вибір персонажів, логіку втрат і винагород, трансформацію героя, дистанцію між автором, наратором і персонажем, а також через те, чия перспектива подається як привілейована або проблематизована (Bal, 2017; Booth, 1983; Chatman, 1978; Rimmon-Kenan, 2002).

Дискурсивний рівень фіксує способи соціальної категоризації, легітимації та розподілу агентності. У логіці Феркло дискурс пов'язує текст із дискурсивною та соціальною практикою (Fairclough, 1992), а підхід ван Дейка дає змогу аналізувати ціннісні ієрархії через стратегії репрезентації «своїх» і «чужих», асиметрії відповідальності та ідеологічне позиціонування (van Dijk, 1998, 2008). Для цього дослідження важливо, що художній текст не лише називає соціальних акторів, а й розподіляє між ними право діяти, говорити, оцінювати, бути виправданими або засудженими.

Інтертекстуальний рівень показує, як цінності активуються через зв'язок конкретного твору з попередньо написаними текстами, культурними прецедентами й символічними авторитетами. Ідея тексту як «мозаїки цитат» (Kristeva, 1980) і концепція гіпертекстуальності Женетта (Genette, 1997b) дозволяють розглядати відсилання до канону, біблійних мотивів, радянського культурного коду, козацького міфу, пам'яті Голодомору, Майдану або війни не лише як стилістичний прийом, а як форму ціннісного позиціонування. У художньому контексті пострадянської України цей рівень набуває особливого значення. Перевизначення культурної ідентичності, зміни в розумінні пам'яті та ієрархій символів активно впливають на способи організації наративів і ціннісних опозицій (Гундорова, 2012).

Окремо слід розрізняти маніфестні й латентні репрезентації. Маніфестні цінності прямо декларуються в оцінних судженнях персонажів, наратора або авторизованої позиції. Латентні цінності не завжди називаються, але діють як нормативний фон тексту - як те, що подається очевидним, природним або таким, що не потребує виправдання. Тому контент-аналіз маніфестних одиниць (Berelson, 1952) має поєднуватися з відтворюваним переходом від текстових даних до контексту їх використання (Krippendorff, 2019) і з якісною інтерпретацією латентних смислів.

Окремим важливим розрізненням, яке буде використано в емпіричній частині дослідження (підрозділи 4.5–4.6), є розрізнення між **режимами репрезентації** цінностей: лексично-маркований режим (експліцитна оцінна лексика, прямі ціннісні концепти, безпосередні номінації) і нелексикалізований режим (наративно-фоновий, сюжетно-композиційний, символічний, інтертекстуальний). Передбачається, що одна й та сама цінність може у різні історичні періоди репрезентуватися переважно в одному з цих режимів. Розходження між зростанням або стабільністю цінності у декларативно-опитувальному режимі (WVS, EVS) і зниженням її лексичних маркерів у художньому корпусі саме по собі не доводить послаблення чи посилення цінності; у цій роботі воно розглядається як підстава для обережної інтерпретації в термінах дисоціації режимів репрезентації — можливого зсуву переважного способу її вираження в художньому дискурсі.

Підсумовуючи оглянуті особливості рівнів репрезентації цінностей можна зазначити що ціннісний зміст художнього тексту має багаторівневу природу. Він реалізується через оцінну лексику, наративну організацію конфліктів і виборів, дискурсивні механізми категоризації та інтертекстуальне залучення культурно авторитетних смислів. У подальшому емпіричному аналізі ці рівні можуть бути операціоналізовані через різні типи індикаторів: частотність, сполучуваність і тональність ціннісних понять; типи конфліктів і траєкторії персонажів; стратегії номінації, предикації, легітимації та розподілу агентності; відсилання до культурних канонів і історичних прецедентів. Таке розмежування створює підґрунтя для інтегрованого текстоцентричного підходу, у якому кількісна фіксація повторюваних маркерів поєднується з якісною інтерпретацією їхнього наративного, дискурсивного та культурного значення.

1.4. Текстоцентричний підхід до аналізу ціннісних орієнтацій у художньому дискурсі

Текстоцентричний підхід у системі методів дослідження цінностей: обґрунтування вибору для корпусного аналізу художнього дискурсу.

Методологічний вибір цього дослідження визначається типом даних і характером дослідницького завдання. Серед класів методів, придатних для вивчення цінностей (опитувальні та психометричні, проєктивні та інтерпретативні, поведінкові та імпліцитні, текстоцентричні — відповідно до класифікації, запропонованої в авторській публікації (Маїлян, 2026)), у роботі обрано текстоцентричний підхід — як такий, що дає змогу працювати з уже існуючими текстовими артефактами тривалого часового охоплення.

У межах цієї роботи текстоцентричний підхід розуміється як дослідницька стратегія, за якої основним емпіричним матеріалом реконструкції ціннісних орієнтацій виступає не самозвіт респондента і не безпосередньо спостережувана поведінка, а структурована сукупність текстів, що аналізується на кількох взаємопов'язаних рівнях — лексико-семантичному, наративному, дискурсивному та інтертекстуальному (детально розглянутих у підрозділі 1.3). Такий підхід не ототожнюється з одним методом, а становить рамку інтеграції кількісних і якісних аналітичних процедур.

На відміну від опитувальних і поведінкових методів, які дають синхронні зрізи декларованих ціннісних пріоритетів або ситуативних кореляцій (Hitlin & Piliavin, 2004; Maio, 2016; Schwartz, 1992, 2012, 2021; Schwartz et al., 2012; World Values Survey Association, n.d.-a), текстоцентричний підхід дозволяє реконструювати ціннісні репрезентації завершених історичних періодів. Ретроспективна реконструкція ціннісних зрушень через опитування є обмеженою для завершених історичних періодів: носіїв тих ціннісних позицій

або вже неможливо опитати, або їхні відповіді знаходились би під впливом пізніших досвідів і переосмислень (Krippendorff, 2019). Художні тексти 1991–2025 рр., натомість, існують як завершений масив даних, незалежний від ретроспективних спотворень пам'яті респондентів. Вони фіксують ціннісні репрезентації у формі, в якій ті були введені в публічний культурний обіг у певний історичний період (McEnery & Hardie, 2012; Grimmer et al., 2022) — попри можливі часові лаги між написанням, публікацією та суспільною рецепцією тексту. У підготовленій до друку авторській статті з методології аналізу ціннісних систем у художньому дискурсі обґрунтовано, що жоден окремий метод не забезпечує повної реконструкції ціннісного змісту тексту; найбільш продуктивним є їх поєднання в межах інтегрованого текстоцентричного підходу (Маїлян, 2026).

Разом із тим текстоцентричний підхід у дослідженні з масштабним корпусом і тривалим часовим охопленням ставить дослідника перед двома взаємно протилежними методологічними викликами. Перший — кількісний: масштаб корпусу унеможливорює повноцінне глибинне прочитання кожного тексту засобами класичного якісного аналізу (McEnery & Hardie, 2012; Grimmer et al., 2022). Другий — якісний: суто автоматизований кількісний підхід не здатен розкрити латентні ціннісні смисли, дискурсивні механізми легітимації / делегітимації ціннісних позицій та наративну логіку ціннісних зрушень (Nelson, 2020). Вирішення цієї суперечності вимагає ієрархічної інтеграції методів: пріоритетного — для систематичного кількісного охоплення корпусу, і допоміжних — для поглибленої якісної інтерпретації відібраних кейсів. Реалізація такої ієрархії передбачає прозоре формування корпусу як цільової вибірки текстових джерел, принципи якого розгорнуто в підрозділі 2.1.

Зазначена методологічна рамка визначає ієрархічну структуру інструментарію дослідження, яка детально описується в наступних підрозділах: контент-аналіз обирається як пріоритетний метод систематичної

кількісної фіксації ціннісних маркерів у межах усього корпусу та кожного хронологічного зрізу (Krippendorff, 2019); КДА, наративний аналіз, сентимент-аналіз та NLP-процедури виконують роль допоміжних процедур контекстуалізації результатів контент-аналізу на матеріалі відібраних кейсів.

Контент-аналіз як пріоритетний метод виявлення ціннісно маркованих одиниць художнього тексту: процедура та аксіологічна адаптація.

Контент-аналіз посідає центральне місце в методологічному інструментарії цього дослідження як пріоритетний метод систематичної роботи з масивом україномовних художніх текстів 1991–2025 рр. На відміну від суто якісного прочитання, контент-аналіз дає змогу поєднати чутливість до смислової структури художнього тексту з відтворюваною процедурою кодування, кількісною агрегацією результатів і подальшою якісною інтерпретацією (Krippendorff, 2019). Його застосування зумовлене потребою не лише інтерпретувати окремі твори, а й забезпечити порівнювану фіксацію повторюваних ціннісних маркерів у межах різних хронологічних зрізів — завдання, для якого опитувальний інструментарій є обмеженим і потребує доповнення текстоцентричними процедурами за причин, обґрунтованих у підрозділі 1.4. Детальні теоретичні засади методу та його місце серед інших текстоцентричних підходів попередньо розглянуто в авторській статті (Маїлян, 2026); у цьому підрозділі акцент зроблено на аксіологічній адаптації методу до завдань магістерського дослідження.

Принциповою для аксіологічного аналізу є та обставина, що контент-аналіз не вимірює ціннісні орієнтації безпосередньо, як це відбувається в опитувальних дослідженнях, а фіксує їхні текстові репрезентації: оцінну лексику, моральні судження, модальні конструкції, символічні образи, сюжетні

конфлікти та повторювані моделі належної або засуджуваної поведінки. Тому одиниці аналізу визначаються на кількох рівнях, що відповідають введеним у підрозділі 1.3 рівням репрезентації цінностей: на лексико-семантичному рівні фіксуються слова й словосполучення з вираженою оцінністю; на наративному — репліки персонажів, внутрішні монологи, конфлікти й сюжетні ситуації, у яких прямо або опосередковано виражено ставлення до певних дій, явищ чи соціальних ролей; на дискурсивному — стійкі моральні опозиції, способи позиціювання соціальних акторів і форми легітимації або проблематизації певних моделей поведінки.

Аксіологічна адаптація контент-аналізу для цього дослідження спирається на чотири принципи. По-перше, відтворюваність процедури вимагає чіткого розрізнення одиниці аналізу, одиниці кодування, одиниці рахунку та одиниці контексту (Krippendorff, 2019; Neuendorf, 2017); операціоналізацію цих одиниць для україномовного художнього корпусу подано в підрозділі 2.2. По-друге, категоріальна система поєднує дедуктивний рівень (на основі теорії базових цінностей Шварца, розглянутої в підрозділі 1.1) та індуктивний рівень, що враховує специфіку українського соціокультурного досвіду 1991–2025 рр. — національну ідентичність, колективну пам'ять, опозицію «свій / чужий», жертвовність, моральну відповідальність; конкретний склад категорій подано в підрозділі 2.2 (Fischer et al., 2022; Schwartz, 2006, 2012, 2021). По-третє, автоматизоване кодування пріоритетно фіксує маніфестний шар (експліцитна оцінна лексика, моральні судження, модальні конструкції), тоді як латентний шар (повторювані символічні структури, моральні опозиції, сюжетні наслідки) виявляється переважно в якісній інтерпретації відібраних кейсів і потребує описаних у кодбуку правил роботи з неоднозначними випадками (Berelson, 1952; Schreier, 2012). По-четверте, надійність кодування перевіряється коефіцієнтом α Криппендорфа (Hayes & Krippendorff, 2007; Krippendorff, 2019), процедуру та результати застосування якого подано в Розділі 3.

Комп'ютерно-асистовані інструменти виконують допоміжну роль у реалізації контент-аналізу — підтримують пошук маркерів, нормалізацію частот і порівняння результатів між хронологічними зрізами (Grimmer et al., 2022; Nelson, 2020); конкретний технологічний стек цього дослідження описано в підрозділі 2.2. Інтерпретативне поглиблення результатів контент-аналізу через допоміжні процедури розглядається в наступному підрозділі.

Допоміжні процедури контекстуалізації результатів контент-аналізу.

Контент-аналіз формує кількісно впорядкований фундамент аналізу ціннісних репрезентацій у корпусі, однак не вичерпує методологічних можливостей текстоцентричного підходу. Три рівні репрезентації цінностей у художньому тексті, розглянуті в підрозділі 1.3 (лексико-семантичний, наративний, дискурсивний), не потребують розгортання окремих повномасштабних процедур у межах усього корпусу, однак вимагають контекстуальної перевірки й інтерпретативного уточнення результатів кількісного контент-аналізу на матеріалі відібраних кейсів. Цю функцію в дослідженні виконують критичний дискурс-аналіз (КДА), наративний аналіз, сентимент-аналіз і NLP-процедури. Інтертекстуальний рівень, виокремлений у підрозділі 1.3, не оформлюється як окрема процедура аналізу всього корпусу, оскільки його автоматизована фіксація потребувала б окремої бази культурних прецедентів і спеціального дизайну дослідження; натомість інтертекстуальні маркери враховуються в якісному аналізі відібраних кейсів у межах КДА та наративної інтерпретації.

Критичний дискурс-аналіз (КДА) застосовується для з'ясування механізмів, через які художній текст позиціонує, легітимує або делегітимує певні ціннісні позиції. Спираючись на трирівневу модель Н. Феркло (текст — дискурсивна практика — соціальна практика) (Fairclough, 1992, 2003), аналіз

кейсів у цій роботі враховує не лише мовну організацію фрагмента, а й способи його включення в ширший соціокультурний контекст періоду. Особливо релевантним для аксіологічного аналізу художнього дискурсу України 1991–2025 рр. є підхід ван Дейка до аналізу бінарних опозицій «свій / чужий» (van Dijk, 1998, 2008), який допомагає інтерпретувати конфліктний вимір категорії UA-IDN. У процедурному відношенні КДА застосовується до цілеспрямовано відібраних фрагментів у межах діагностичної якісної перевірки показових і контрастних кейсів, розгорнутої в підрозділі 4.4 (Reisigl & Wodak, 2015; Titscher et al., 2000).

Наративний аналіз охоплює рівень репрезентації цінностей, що виявляється через організацію подій, моральний вибір персонажів у кризових ситуаціях, дистанцію між наратором, персонажем і передбачуваною авторською позицією, а також оцінний вимір наративу — те, що в тексті постає як значуще, нормативне або таке, що порушує очікуваний нормативний порядок (Bal, 2017; Chatman, 1978; Labov & Waletzky, 1967; Riessman, 2008; Rimmon-Kenan, 2002). Як і КДА, наративний аналіз застосовується не до всього корпусу, а до теоретично обґрунтованого добору показових і контрастних фрагментів у межах діагностичної якісної перевірки, описаної в підрозділі 4.4 (Grimmer et al., 2022; Nelson, 2020); обираються твори або фрагменти, у яких попереднє кількісне кодування виявило найвищі концентрації певних ціннісних категорій, різкі відмінності між хронологічними зрізами або методологічно важливі неоднозначні випадки.

Сентимент-аналіз використовується не як метод виявлення цінностей, а як допоміжна процедура оцінювання емоційно-оцінної тональності лексичного оточення вже зафіксованих ціннісно маркованих одиниць у словнику Tone-Dict-Ukrainian (lang-uk, n.d.; Pang & Lee, 2008). Принципово важливе застереження: тональність не дорівнює ціннісному знаку і не є самостійною підставою для присвоєння або зняття категорії, а залучається лише як описовий параметр для

якісного аналізу. Технічну реалізацію (словник Tone-Dict-Ukrainian, вікно ± 5 токенів) описано в підрозділі 2.2. Через відомі обмеження сентимент-моделей для україномовного художнього тексту — особливо щодо іронії, наративної перспективи та контекстуальної переоцінки тональності (Booth, 1983; Liu, 2012; Shynkarov et al., 2025) — результати інтерпретуються виключно як допоміжні індикатори, що підлягають перевірці в межах КДА і наративного аналізу.

NLP-процедури загального призначення (лематизація, частотний аналіз, побудова контекстних вікон) застосовуються на корпусному рівні для технічного забезпечення контент-аналізу; конкретні інструменти описано в підрозділі 2.2 (McEnery & Hardie, 2012).

Таким чином, чотири допоміжні процедури виконують функціонально диференційовані завдання: КДА використовується для контекстуального уточнення дискурсивного рівня репрезентації цінностей, наративний аналіз — для інтерпретації окремих сюжетно-ціннісних конфігурацій, а сентимент-аналіз і NLP-процедури — для технічної підтримки лексико-семантичного та корпусного рівнів. Принципи поєднання всіх методів, включно з процедурою хронологічного порівняння та методологічними обмеженнями, розглядаються в наступному підрозділі.

Інтеграція методів, порівнянність хронологічних зрізів і методологічні обмеження текстоцентричного аналізу.

Інтеграція методів у цьому дослідженні має функціональний, а не адитивний характер: автоматизований контент-аналіз є базовим методом, що забезпечує масштабовану й відтворювану фіксацію ціннісно маркованих одиниць у корпусі; статистичні процедури і жанрова стандартизація опрацьовують агрегатну динаміку та коригують ефект нерівномірної жанрової композиції між зрізами; процедури валідації (Розділ 3) визначають межі довіри

до автоматизованого кодування; КДА і наративний аналіз виконують діагностичну інтерпретацію відібраних кейсів (підрозділ 4.4); сентимент-аналіз і NLP-процедури забезпечують технічну та описову підтримку основного методу. Така ієрархія фіксує субординацію процедур: висновки роботи спираються насамперед на результати контент-аналізу з валідацією, а допоміжні процедури служать їхньому контекстуальному уточненню.

Порівнянність хронологічних зрізів забезпечується єдиною процедурою підготовки текстів, спільною категоріальною системою і використанням відносних показників активності категорій (підрозділи 2.1–2.2); вона не є абсолютною через відмінності корпусу за обсягом, жанровою структурою і цифровою доступністю творів. Тому результати аналізу інтерпретуються не як пряме вимірювання цінностей українського суспільства, а як динаміка словниково зафіксованих маркерів ціннісних репрезентацій у межах сформованого корпусу. Художній текст є естетично, жанрово й інституційно опосередкованим джерелом, а автоматизоване кодування не завжди здатне розрізнити іронію, фокалізацію, позицію персонажа й наратора або латентний ціннісний конфлікт. Кількісні результати тому розглядаються як підстава для інтерпретації, а не самодостатній доказ трансформації ціннісних орієнтацій. Така логіка безпосередньо переходить до формулювання теоретично мотивованих очікувань, які перевіряються в емпіричній частині дослідження.

1.5. Теоретично мотивовані очікування про динаміку ціннісних репрезентацій

Концептуальна рамка і операціоналізована періодизація (підрозділ 1.2) дозволяють сформулювати три теоретично мотивовані очікування про напрям руху ціннісних репрезентацій в україномовній художній прозі 1991–2025 рр. Ці

очікування формуються у двох комплементарних формах: як додаткові діагностичні очікування на агрегатному рівні корпусу (H1, H2, H3) і як уточнюючі стратифіковані очікування (H1s, H2s, H3s), що враховують жанрову неоднорідність корпусу.

Очікування 1 (H1 / H1s). Активність категорії UA-IDN (національна ідентичність) має демонструвати локальні підвищення в періодах, що йдуть слідом за ключовими громадянсько-мобілізаційними подіями: 2004–2013 рр. (після Помаранчевої революції) і 2014–2021 рр. (після Революції Гідності та початку війни на Донбасі). H1s: ця динаміка може бути виразнішою у жанрових блоках, орієнтованих на сучасну тематику (А, В), і слабшою у блоці Е, де переважає історична проза з власною ретроспективною ідентичнісною лексикою.

Очікування 2 (H2 / H2s). Активність категорії ENH (самоствердження) має демонструвати тенденцію до зростання впродовж досліджуваного періоду, що відповідає постсоціалістичному модернізаційному зсуву від колективістських до індивідуалістичних орієнтацій у термінах моделі Інглхарта-Велзеля (Inglehart & Welzel, 2005; World Values Survey Association, n.d.-a). H2s: цей зсув очікувано виразніший у жанровій белетристиці (блок В) і дитячо-підлітковій прозі (блок D), де представлені сучасні моделі особистого успіху, кар'єри, стилю життя.

Очікування 3 (H3 / H3s). Активність категорії TRA (самотрансцендентність) має демонструвати посилення у періоді 2022–2025 рр. як реакція на воєнний контекст — зростання мотивів солідарності, жертвовності, турботи про близьких і колективної підтримки. H3s: цей сплеск очікувано фіксується насамперед у жанрах, безпосередньо орієнтованих на сучасний воєнний досвід (блок А, частково В), і менш виразний у блоці Е (історична проза) і блоці С (фантастика і фентезі).

Методологічне очікування (Н4, додатково). Жанрова структура корпусу істотно змінюється між хронологічними зрізами (підрозділ 4.2). З огляду на це сирі агрегатні показники окремих категорій можуть відображати не саму ціннісну динаміку, а зміни жанрового складу корпусу. Пряма стандартизація за фіксованою жанровою структурою (підрозділ 4.2) має суттєво модифікувати інтерпретацію сирих трендів — у деяких випадках до зміни їхнього напрямку.

Емпіричну перевірку Н1–Н3 і Н1s–Н3s, а також аналіз методологічного очікування Н4, наведено у підрозділах 4.3 і 4.2.

1.6. Висновки до розділу 1

У Розділі 1 сформовано теоретико-методологічну основу дослідження трансформації ціннісних орієнтацій у дискурсі україномовної художньої прози 1991–2025 рр. Розмежовано чотири категорії аксіологічного аналізу: цінність як надситуативний соціокультурний орієнтир оцінки і дії; ціннісну орієнтацію як спосіб її привласнення суб'єктом чи групою; ціннісну систему як структуровану конфігурацію взаємопов'язаних цінностей; ціннісну установку як ситуативну готовність до оцінки чи дії. Прийнято соціологічний ракурс розуміння цінностей як елементів колективних нормативних систем, культурних класифікацій і соціальних практик, що уможливлює використання художньої прози як легітимного емпіричного джерела. Класична традиція (Durkheim, 1912/1995; Parsons, 1951; Thomas & Znaniecki, 1918; Weber, 1946a) задає інституційно-нормативний рівень аналізу, а сучасні моделі (Bourdieu, 1990; Inglehart, 1977; Schwartz, 1992) забезпечують операціоналізацію. Теорія базових цінностей Шварца використовується як концептуальне підґрунтя категоріальної системи дослідження.

Соціальну трансформацію концептуалізовано як рамку аналізу ціннісних змін із трьома аналітичними типами — кризою, зрушенням і реконструкцією. Український випадок 1991–2025 рр. описано як нелінійну траєкторію з чотирма хронологічними зрізами: 1991–2003 рр. (пострадянська аномія, переважання цінностей виживання); 2004–2013 рр. (ціннісна поляризація і хвилі громадянської мобілізації); 2014–2021 рр. (прискорене перевизначення ідентичностей в умовах війни на Донбасі); 2022–2025 рр. (ціннісна реконструкція в умовах повномасштабного вторгнення). Ця періодизація виконує роль операціоналізованої моделі соціального контексту, що задає інтерпретаційну та порівняльну рамку для подальшого аналізу.

Послідовно розмежовано художній текст і художній дискурс: текст є одиницею відбору і кодування ціннісних маркерів, дискурс — рівнем їхньої соціологічної контекстуалізації в межах літературного поля та фаз соціальної трансформації. Художню прозу використано як соціологічно інформативне, але не самодостатнє джерело; виокремлено чотири рівні репрезентації цінностей — лексико-семантичний, наративний, дискурсивний та інтертекстуальний.

Методологічну рамку сформовано як ієрархічно інтегрований текстоцентричний підхід. Автоматизований аксіологічний контент-аналіз обрано як пріоритетний метод систематичної фіксації ціннісних маркерів у корпусі та хронологічних зрізах; критичний дискурс-аналіз, наративний аналіз, сентимент-аналіз і NLP-процедури виконують допоміжні функції контекстуалізації результатів. Корпус розуміється не як лінгвістична емпірична база, а як цільова соціологічна вибірка текстових артефактів певного історичного періоду. Сформульовано три теоретично мотивовані очікування щодо динаміки ціннісних категорій UA-IDN, ENH і TRA (H1–H3) та одне методологічне очікування щодо ефекту жанрової композиції корпусу (H4), з відповідними стратифікованими варіантами (H1s–H3s), що підлягають емпіричній перевірці в Розділі 4.

Тим самим Розділ 1 окреслює режим допустимих висновків у наступних розділах: ідеться про реконструкцію динаміки текстуально зафіксованих ціннісних репрезентацій у корпусі, а не про пряме вимірювання цінностей українського суспільства. Розділ 2 розгортає цю рамку в емпіричну й інструментальну процедуру, Розділ 3 встановлює межі довіри до результатів через процедури валідації, Розділ 4 здійснює змістовну інтерпретацію виявлених трендів.

РОЗДІЛ 2. КОРПУС І ІНСТРУМЕНТАРІЙ

У цьому розділі описано емпіричну й інструментальну основу дослідження. Розділ структурно ділиться на дві частини: формування й підготовка корпусу україномовної художньої прози 1991–2025 рр. (підрозділ 2.1) та інструментарій автоматизованого аксіологічного контент-аналізу (підрозділ 2.2). Логіка викладу йде від матеріалу дослідження (твори, жанри, періодизація, технічна підготовка) до процедури його аналітичної обробки (одиниці кодування, категоріальна система, технологічний стек, валідаційні процедури, межі підходу). Особливу увагу приділено фіксації обмежень — як на рівні самого корпусу (підрозділ 2.1), так і на рівні автоматизованого підходу (підрозділ 2.2).

2.1. Формування та підготовка корпусу текстів дослідження

Поняття корпусу в цій роботі використовується у міждисциплінарному, але передусім соціологічному значенні. Якщо в лінгвістиці корпус зазвичай позначає структуровану сукупність текстів, призначену для аналізу мовних явищ, то в межах цього дослідження він трактується як емпірично організований масив культурних артефактів. Його соціологічна функція полягає не у вимірюванні індивідуальних установок авторів або читачів, а у фіксації тих способів репрезентації цінностей, які були введені в публічний художній обіг у певні історичні періоди. Емпіричну основу дослідження становить корпус україномовних художніх прозових творів 1991–2025 рр. Він не є статистично репрезентативною вибіркою всієї україномовної художньої прози 1991–2025 рр., а виконує функцію цільового аналітичного масиву текстових джерел, через який

реконструюються ціннісні репрезентації, що циркулювали у публічному художньому дискурсі в межах чотирьох хронологічних зрізів. Процедура формування корпусу описана нижче за послідовною логікою: критерії включення >> джерела >> жанрова класифікація >> періодизація >> технічна підготовка >> побудова аналітичного датасету >> фіксація обмежень. Кожен крок документовано відповідно до вимог наукового контент-аналізу (Krippendorff, 2019).

Критерії відбору творів до корпусу.

Відбір творів ґрунтувався на системі критеріїв включення та виключення, що дозволяють сформувати цільовий, тематично релевантний і мовно однорідний аналітичний корпус україномовної художньої прози, придатний для дослідження ціннісних репрезентацій у межах визначених хронологічних зрізів.

До критеріїв включення належать: перша публікація твору в часовому діапазоні від 1991 по 2025 рік включно; написання твору безпосередньо українською мовою без перекладу з інших мов; наявність оцифрованого тексту в одному з прийнятних форматів (PDF, DJVU, EPUB, MOBI, FB2, HTML, HTM, DOC, DOCX, RTF, TXT); належність до художньої прозової форми з перевагою творів великого та середнього обсягу (романів, повістей, збірок новел та оповідань). Якщо твір мав кілька форм появи у публічному обігу, пріоритет надавався першій книжковій або авторитетно зафіксованій публікації, яка дозволяла віднести текст до відповідного хронологічного зрізу. Повторні перевидання без істотних текстуальних змін не розглядалися як нові одиниці корпусу. Двомовні автори (зокрема ті, хто видає твори і українською, і російською мовою) включаються до корпусу лише за тими творами, перша публікація яких була україномовною; перекладні та автоперекладні твори виключаються незалежно від статусу автора.

Поняття "корпус" у межах цього дослідження використовується не у вузько-лінгвістичному значенні (як емпірична база для опису мови), а в соціологічному: під корпусом ми розуміємо цільову, теоретично обґрунтовану вибірку текстових артефактів — україномовних художніх прозових творів 1991–2025 рр., — структуровану за критеріями жанру, часу публікації та доступності, що використовується як емпірична база для реконструкції суспільно значущих ціннісних репрезентацій певного історичного періоду. Такий підхід відповідає традиції соціології культури (Wuthnow, 1987), культурно-соціологічного аналізу літератури (Váňa, 2020) та «дистантного читання» як методу великомасштабного текстового аналізу культурних процесів (Boyd et al., 2022; Moretti, 2005).

До критеріїв виключення належать: повторні перевидання без значних текстуальних змін; переклади з інших мов на українську, зокрема випадки, коли автор спочатку публікував твір іншою мовою і лише згодом здійснював переклад самостійно або із залученням перекладачів; поетичні, віршовані та драматичні твори; наукова, публіцистична, журналістська, довідкова та суто документальна продукція, яка не має художньо-нарративної організації; антології та збірники оповідань кількох авторів; казкова та розважальна дитяча проза для молодшого віку, якщо вона не має достатнього обсягу, авторської художньої структури та релевантності для аналізу ціннісних репрезентацій. Художньо-документальна проза включалася до корпусу лише за умови наявності виразної літературно-нарративної організації: сюжетної побудови, авторської композиції, образної системи або художньо опрацьованої репрезентації досвіду. Суто інформаційні, журналістські, довідкові, науково-популярні або документальні тексти без художньої організації до корпусу не включалися.

Виняток із цього критерію виключення складає підліткова проза та *Young Adult*: вони розглядалися як окремий значущий сегмент корпусу, оскільки репрезентують моделі ціннісного самовизначення, соціалізації та ідентичності,

релевантні для аналізу. Деталізована логіка віднесення дитячих творів до корпусу — у Додатку В (блок D).

Поетичні твори не включалися до корпусу, оскільки вони потребують іншої одиниці аналізу, іншої логіки сегментації тексту та окремих процедур інтерпретації ритміко-синтаксичної, строфічної й образної організації. З огляду на те, що в цьому дослідженні базовою одиницею запису є абзац прозового тексту, включення поезії порушувало б методологічну однорідність корпусу.

Джерела пошуку творів і формування списку.

Першочергово до пошуку включалися твори авторів, які отримували літературні відзнаки або потрапляли до авторитетних рекомендаційних переліків. Основними орієнтирами були публічні переліки лауреатів і номінантів Національної премії України імені Тараса Шевченка, PEN Ukraine / «Сенсор» («100 книжок, які допоможуть зрозуміти Україну»), конкурсу «Коронація слова», «Книги року BBC» і конкурсу імені Григора Тютюнника (BBC News Україна, n.d.; PEN Ukraine & Сенсор, n.d.; Комітет з Національної премії України імені Тараса Шевченка, n.d.; Конкурс імені Григора Тютюнника, n.d.; Коронація слова, n.d.).

Як допоміжні джерела використовувалися також окремі регіональні та всеукраїнські літературні конкурси за умови наявності публічного реєстру номінантів або переможців, повторюваності конкурсу та його релевантності саме для художньої прози. Ці джерела мали допоміжний, а не визначальний статус і використовувалися переважно для розширення первинного списку та перевірки літературної видимості окремих авторів.

Якщо один і той самий твір фігурував у кількох джерелах, він включався до корпусу один раз, а інформація про кілька джерел фіксувалася в метаданих. За наявності кількох цифрових копій перевага надавалася файлу з кращою

структурою, повнішими метаданими, меншою кількістю OCR-артефактів і придатністю до автоматизованої обробки.

Додатковим бібліографічним джерелом для верифікації та розширення переліку слугував портал Книжкової палати України імені Івана Федорова та його бібліографічні літописи (Книжкова палата України імені Івана Федорова, n.d.).

Первинний список творів, складений за описаними джерелами, налічував близько 2 600 позицій. У процесі формування фінального корпусу позиції відсіювалися у такому порядку: недоступність цифрової копії в задовільному форматі; дублювання (один твір у кількох джерелах); уточнення статусу твору як перекладного або двомовно-перекладного; невідповідність критеріям виключення (поезія, нехудожні жанри, повторні перевидання). Документація поетапного відбору зберігається в системі логів обробки корпусу. До фінальної робочої версії корпусу включено 664 твори 235 авторів. Залишкові технічні обмеження описано в підрозділі 2.1.

Синтетична класифікація жанрів творів корпусу

Для систематизації корпусу розроблено синтетичну класифікацію, яка диференціює твори за жанровою ознакою. Синтетичний характер класифікації полягає в тому, що вона поєднує літературознавчі жанрові ознаки, видавничо-рецептивні категорії та дослідницьку релевантність жанру для аналізу ціннісних репрезентацій.

Жанрова класифікація передбачає 18 основних категорій (G1–G18), згрупованих у п'ять тематичних блоків (Таблиця 2.1; повний кодифікований опис — у Додатку В).

Таблиця 2.1.

Жанрові блоки і коди корпусу (зведений вигляд)

Блок	Тематичний фокус	G-коди
A	Реалістична та сучасна проза	G1 (соціальна), G2 (психологічна), G3 (автобіографічна, мемуарна), G4 (воєнна), G5 (міська)
B	Жанрова белетристика	G6 (детектив і піджанри), G7 (жіноча проза), G8 (романтична), G9 (пригодницька)
C	Фантастика і фентезі	G10 (наукова фантастика і піджанри), G11 (фентезі та піджанри), G12 (містика та горор)
D	Дитяча та підліткова проза	G13 (дитяча), G14 (підліткова, YA)
E	Гібридні та пограничні форми	G15 (постмодерна), G16 (художньо-документальна, non-fiction), G17 (історична), G18 (гібридні жанри)

Примітка. Категорія G7 («жіноча проза») використовується як умовна літературно-видавнича й аналітична категорія за вживанням у видавничій рецепції 1990-х — 2010-х рр., а не як есенціалістське позначення «жіночого письма» в гендерному сенсі.

Жанрова класифікація ґрунтується на засадах генологічного підходу Бовсунівської (Бовсунівська, 2009, 2010) з урахуванням специфіки сучасної української літератури та її жанрових гібридів (Качак, 2017-2018). На етапі попереднього маркування використано велику мовну модель Claude Opus у режимі few-shot класифікації за операційними описами 18 категорій; усі 664 присвоєння жанрових кодів пройшли суцільну ручну верифікацію автора (деталі двоступеневої процедури — у Додатку В) (Gilardi et al., 2023; Grimmer et al., 2022; Ziems et al., 2024).

Жанрова класифікація використовується не лише як описова характеристика корпусу, а й як аналітична змінна. Вона дозволяє перевіряти, чи відрізняються ціннісні репрезентації залежно від жанрового блоку: наприклад, чи домінують у воєнній прозі цінності безпеки, жертвності й солідарності, у детективі — справедливості та інституційного порядку, а в підлітковій прозі — самовизначення й автономії. У ширшому методологічному сенсі така класифікація відповідає практикам побудови літературно-історичних корпусів,

де метадані про жанр, час публікації та тип тексту є необхідною умовою подальшого порівняльного аналізу (Fileva, 2023).

Для подальшої інтерпретації результатів аксіологічного контент-аналізу принципово важливо зафіксувати, що жанрова композиція корпусу істотно змінюється від раннього зрізу до пізнього (Таблиця 2.2). Найбільш виражений зсув — між блоком Е (гібридні та пограничні форми, що домінують у 1991–2003 рр. через високу частку історичної прози G17) і блоком С (фантастика і фентезі, який зростає у понад вісім разів між першим і четвертим зрізом).

Таблиця 2.2.

Жанрова композиція корпусу за хронологічними зрізами, % абзаців

Блок	1991–2003	2004–2013	2014–2021	2022–2025	Δ (п.п.)	Δ (%)
А	27,26	30,8	22,0	18,97	–8,29	–30,4
В	12,18	19,82	18,43	23,67	11,49	94,3
С	4,04	9,65	16,14	35,19	31,15	770,7
Д	3,17	11,14	7,61	6,49	3,31	104,4
Е	53,35	28,59	35,82	15,69	–37,66	–70,6
Разом	100	100	100	100	—	—

Примітка. А — реалістична та сучасна проза; В — жанрова белетристика; С — фантастика і фентезі; Д — дитяча та підліткова проза; Е — гібридні та пограничні форми, у складі яких домінує історична проза G17.

Формування часової періодизації творів.

Між завершенням рукопису і публікацією існує часовий лаг, тому рік видання не є точним індикатором моменту формування ціннісного змісту твору. У цьому дослідженні він використовується як маркер входження тексту в публічний літературний обіг — момент видавничого рішення, готовності читачів і критики до сприйняття певних позицій і (з певним лагом) завершення

авторської роботи. Така операціоналізація узгоджується з логікою кількісної літературної історії Ф. Моретті (Moretti, 2005).

На підставі цього підходу виділено чотири хронологічні зрізи за роками видання: **1991–2003** — період пострадянської дезорієнтації до Помаранчевої революції; **2004–2013** — період між Помаранчевою революцією і Революцією Гідності; **2014–2021** — період між початком російської агресії на Донбасі та повномасштабним вторгненням; **2022–2025** — період після 24 лютого 2022 р. Така стратифікація спирається на соціально-історичну рамку, обґрунтовану в підрозділі 1.2, і відповідає загальній логіці корпусного аналізу, у якому часові підкорпуси дають змогу порівнювати динаміку мовних і змістових показників у різних періоди (McEnery & Hardie, 2012).

Останній зріз 2022–2025 рр. має особливий статус: він є найкоротшим за тривалістю, охоплює незавершений літературний процес повномасштабної війни та залежить від обмеженої цифрової доступності найновіших текстів. Тому його результати інтерпретуються не як стабілізований підсумок нової літературної фази, а як попередній корпусний зріз ранньої воєнної та післяшокової репрезентації ціннісних категорій.

Твори, видані в перший рік кожного зрізу, інтерпретуються з урахуванням можливого лагу між написанням і публікацією. Тому об'єктом аналізу є не миттєвий ціннісний стан суспільства, а репрезентації, введені в публічний літературний обіг відповідного періоду.

Технічна підготовка текстових файлів.

Знайдені файли надходили в різних форматах і потребували уніфікації. Для PDF- і DJVU-файлів без машиночитаного шару застосовувалося OCR (комерційне рішення з підтримкою української мови); для решти форматів — пряме конвертування зі збереженням структури тексту. Усі файли приведені до

єдиної ієрархічної JSON-схеми (книга - розділи - абзаци) з фіксацією релевантних метаданих (автор, назва, рік, мова, видавець, жанр) і характеристик кожного абзацу (тип, ознака діалогу, кількість слів); демонстративний приклад структури такого JSON-файлу твору з основними полями і вкладеною ієрархією розділів і абзаців наведено в Додатку Г. Контроль якості здійснювався автоматичним скриптом валідації за 27 параметрами, згрупованими у п'ять блоків (мовна ідентичність, повнота метаданих, структурна цілісність, якість тексту, узгодженість аналітичних полів); деталі процедур, переліки кодових позицій та критерії валідації наведено в Додатку Д.

Побудова аналітичного датасету.

Остаточним етапом підготовки є денормалізація JSON-файлів у плоску аналітичну таблицю, де кожен рядок відповідає одному абзацу художнього тексту з повним набором метаданих батьківського документа. Абзац обрано як базову одиницю запису, оскільки він однозначно ідентифікований, відтворювано виділюваний і достатньо малий для точної локалізації маркера, але достатній для утримання ціннісного смислу (Krippendorff, 2019). У ієрархії одиниць твір функціонує як одиниця відбору корпусу, абзац — як одиниця запису, а ширший фрагмент (розділ, сцена) — як контекстна одиниця якісної інтерпретації.

При побудові датасету застосовуються два фільтри: включаються лише абзаци типу prose (виключаються цитати, датування та інші паратекстуальні утворення в термінах Женетта (Genette, 1997b)); виключаються абзаци обсягом менше 5 слів — поріг, узгоджений із загальною вимогою щодо змістовної осмисленості одиниці запису (Krippendorff, 2019; McEnery & Hardie, 2012). Короткі репліки (1–4 слова) можуть враховуватися на етапі якісної інтерпретації ширших контекстних одиниць (підрозділ 2.2).

З огляду на бахтінський принцип розрізнення авторського голосу і голосів персонажів (Bakhtin, 1981, 1984), датасет автоматично розбивається на два субкорпуси: недіалоговий наративний (`is_dialogue = false`) і мовлення персонажів (`is_dialogue = true`). Технічне позначення «авторський наратив» не означає прямого ототожнення недіалогових фрагментів з авторською позицією і потребує якісної інтерпретації на рівні фокалізації, іронії та невластне-прямого мовлення.

Датасет структуровано у чотири групи полів — ідентифікатори, метадані книги, жанрова класифікація, характеристики абзацу; повний опис полів з прикладами значень наведено в Додатку Ж, перелік вторинних файлів датасету — в Додатку Д. Унікальним ключем для міжкнижкового аналізу виступає комбінація ідентифікатора твору і локального ідентифікатора абзацу.

На основі датасету формується окремий файл результатів автоматизованого кодування з бінарними індикаторами семи ціннісних категорій; він використовується для валідації в Розділі 3 і для аналізу корпусних трендів у Розділі 4. Загальну характеристику корпусу подано у Таблиці 2.3; розподіл за хронологічними зрізами — у Таблиці 2.4.

Таблиця 2.3.

Загальна характеристика робочого корпусу

Показник	Значення
Кількість творів	664
Кількість авторів	235
Загальна кількість слів	44 837 606
Загальна кількість абзаців	1 375 333
Недіалоговий наративний субкорпус	691 966 абзаців (50,3%)
Діалоговий субкорпус	683 367 абзаців (49,7%)

Примітка. Розраховано автором на основі файлів `corpus_stats.csv` та `corpus_flat.csv`.

Таблиця 2.4.

Розподіл корпусу за хронологічними зрізами

Період	Книг	Абзаців	Слів
1991–2003	135	234 972	9 411 907
2004–2013	224	408 358	13 076 371
2014–2021	231	514 916	16 141 096
2022–2025	74	217 087	6 208 232
Разом	664	1 375 333	44 837 606

Примітка. Розраховано автором на основі файлів corpus_stats.csv та corpus_flat.csv.

Додаткові показники лексичного обсягу, токенів, унікальних лем і співвідношення наративного та діалогового субкорпусів наведено в Додатку Д; демонстраційний приклад перших рядків таблиці corpus_stats.csv з реальними метаданими творів корпусу подано в Додатку Е для ілюстрації типового формату агрегованих даних на рівні твору. У подальшій інтерпретації враховується підвищена токенна вага блоку Е та різна середня довжина наративних і діалогових абзаців. Таким чином, підготовлений аналітичний датасет являє собою структуровану, верифіковану й методологічно обґрунтовану основу для подальшого кількісного аксіологічного контент-аналізу, допоміжного тонального опису та діагностичної якісної контекстуалізації результатів на матеріалі україномовної художньої прози 1991–2025 рр. (Fileva, 2023; Krippendorff, 2019; McEnery & Hardie, 2012).

Обмеження сформованого корпусу.

Сформований корпус не є статистично репрезентативною вибіркою україномовної художньої прозової літератури 1991–2025 рр., а являє собою цільовий аналітичний масив текстів, які набули літературної, видавничої або преміальної видимості (Biber, 1993). Його обмеження групуються у двох

площинах — методологічній і технічній — і враховуються під час інтерпретації результатів у Розділі 4.

Методологічні обмеження.

Інституційний фільтр преміального відбору. Близько 80% корпусу становлять твори з преміальних або рекомендаційних переліків (детальніше — 2.1), що автоматично виключає самвидавну, виключно цифрову (Wattpad-подібну) і малотиражну прозу, яка не пройшла видавничо-критичних фільтрів. Корпус відображає ціннісні репрезентації «видимої» частини літературного поля, а не всього літературного виробництва (Bourdieu, 1993; English, 2005).

Жанровий перекіс на користь блоку Е. Історична проза G17 становить 21,2% корпусу, тому значна частина ціннісних репрезентацій має ретроспективний характер — конструюється через апеляцію до минулого, а не пряму рефлексію поточного досвіду. Цей перекіс ураховано в Розділі 4 через жанрову стандартизацію (підрозділ 4.2).

Часовий перекіс четвертого зрізу. Зріз 2022–2025 рр. містить лише 74 твори (11,1% корпусу) проти 224 у 2004–2013 рр. і 231 у 2014–2021 рр. — наслідок короткої тривалості періоду (3,5 роки замість 8–10), незавершеності літературного процесу й обмеженої цифрової доступності найновіших текстів. Результати для цього зрізу інтерпретуються з відповідним епістемічним застереженням.

Можливі географічні та гендерні зміщення. Корпус відображає видавничу структуру України з центрами в Києві, Львові та (до 2014 р.) Харкові; регіональна проза з малотиражних видавництв представлена обмежено. Можливі гендерні зміщення у складі авторів не аналізуються як окрема змінна, однак враховуються як потенційна межа інтерпретації.

Технічні обмеження.

OCR-артефакти. 17,9% файлів (119 з 664) містять артефакти автоматичного розпізнавання — злиття виноскових цифр зі словами, бібліотечні штампи в тексті, поодинокі символи-замінники. Артефакти локалізовані переважно в паратекстових і службових абзацах; їхній вплив на основний масив художнього тексту оцінюється як обмежений.

Аномально довгі абзаци. 21,8% файлів (145 з 664) містять принаймні один абзац понад 500 слів — наслідок злиття абзаців при парсингу OCR-файлів. Такі абзаци не вилучалися автоматично, оскільки могли містити цілісний наративний фрагмент, але враховуються як потенційне джерело переважування окремих текстових сегментів; у наступних версіях інструменту доцільне запровадження окремої процедури повторної сегментації.

Відсутність розділової структури. 7,5% корпусу (50 творів) не мають внутрішнього поділу на розділи — текст представлений як єдиний нерозподілений масив. Це обмежує структурний наративний аналіз для цих творів, але не перешкоджає їх включенню до загального кількісного аналізу.

Усі зазначені обмеження задокументовані в системі логів і зафіксовані у відповідних полях якості (quality.warnings) JSON-файлів корпусу.

2.2. Інструментарій автоматизованого аксіологічного контент-аналізу

У цьому підрозділі описано інструментарій, що перетворює корпус, сформований у підрозділі 2.1, на структуровані дані про появу ціннісних маркерів. Методологічну основу становить текстоцентричний підхід, обґрунтований в авторській статті (Маїлян, 2026). Опис охоплює п'ять компонентів: операційні одиниці аналізу і кодування, категоріальну систему із

семи ціннісних категорій, технологічну реалізацію автоматизованого кодування, гібридну процедуру валідації та експліцитну фіксацію меж застосованого підходу. Повний кодбук — операційні визначення категорій, словники маркерів, правила розмежування і каталог граничних випадків — наведений у Додатку И; в основному тексті описано принципи та архітектурні рішення без дублювання деталей реалізації.

Одиниці аналізу і кодування.

Базові одиниці відбору та запису для всього дослідження зафіксовано в підрозділі 2.1: **одиницею відбору** є окремий художній прозовий твір, **одиницею запису** — абзац (paragraph) типу prose з мінімальною довжиною у п'ять слів. У цьому підрозділі ці дві базові одиниці доповнюються трьома категоріями, специфічними для процедури автоматизованого аксіологічного аналізу: одиницею кодування, одиницею рахунку і контекстною одиницею інтерпретації. Сукупність п'яти типів одиниць (Таблиця 2.5) забезпечує методологічно послідовну сходинку від цілісного твору до конкретного маркера ціннісної категорії і назад до інтерпретації в межах ширшого тексту.

Таблиця 2.5.

Система одиниць аналізу і кодування

Тип одиниці	Що це	Звідки	Використання
Одиниця відбору	Окремий художній прозовий твір	Підрозділ 2.1, критерії включення	Формування корпусу 664 творів
Одиниця запису	Абзац типу prose довжиною ≥ 5 слів	Підрозділ 2.1	Базовий рядок аналітичного датасету
Одиниця кодування	Лема або n-грама зі словника категорії	Підрозділ 2.2, кодбук	Автоматичне присвоєння категорії абзацу
Одиниця рахунку	Поява категорії в межах одного абзацу	Підрозділ 2.2	Обчислення частот; нормалізація на 10 000 слів

Контекстна одиниця інтерпретації	Вікно ± 5 токенів (авто) або 1–2 сусідні абзаци (якісно)	Підрозділ 2.2	Контекстні тригери; якісний аналіз кейсів
--	--	---------------	--

Одиниця кодування — це окремий текстовий маркер у межах абзацу, що сигналізує про присутність однієї з категорій кодбуку. У термінах автоматизованого підходу одиницею кодування є лема або n-грама зі словника відповідної категорії. Леми у словнику стратифіковано за трьома класами надійності: strong-леми спрацьовують самостійно як високоспецифічні маркери; standard-леми спрацьовують самостійно з нижчою вагою; weak-леми активують категорію лише за наявності контекстного підкріплення. Ця стратифікація дозволяє включати у словник лексично багатозначні слова (на кшталт «свобода», «перемога», «Україна») без надмірного шуму у кодуванні.

Одиниця рахунку — це зафіксований випадок появи категорії в межах одного абзацу. На кожен абзац може припадати від нуля до семи одиниць рахунку (по одній на кожену категорію). Множинне кодування одного абзацу декількома категоріями є нормою, а не винятком: типовий випадок — фрагмент із пізнішого зрізу, що описує побутову сцену евакуації родини у воєнних умовах, отримує коди TRA (турбота про близьких, доброзичливість), CON (захист родини, безпека) і UA-WAR (контекст війни і переміщення) одночасно. Така структура одиниці рахунку обрана свідомо: вона відображає реальну зв'язаність ціннісних утворень у художньому дискурсі і дозволяє аналізувати не тільки частоти окремих категорій, а й матрицю їхніх спільних появ. Базовим показником для аналізу трендів є частка абзаців, у яких активовано відповідну категорію, у межах певного періоду, жанрового блоку або субкорпусу. За потреби додатково може застосовуватися нормалізація на 10 000 слів, однак основні результати Розділу 4 подаються як відсоток активних абзаців.

Контекстна одиниця інтерпретації використовується на двох рівнях. На рівні автоматичного кодування — це вікно у ± 5 токенів від маркера, в межах

якого скрипт перевіряє наявність контекстних тригерів (наприклад, для розрізнення індивідуальної свободи від колективної національної). На рівні якісного аналізу — це 1–2 сусідні абзаци, які залучаються для оцінки іронії, невластивого мовлення, неоднозначності суб'єкта оцінки. Ширша контекстна одиниця застосовується лише до тих кейсів, які відбираються для якісної інтерпретації; масовий автоматизований прогон використовує тільки вікно у ± 5 токенів.

У повній сукупності одиниць кодувальний інструмент стає інтерпретованим: кожен зафіксований випадок появи категорії можна простежити від конкретного маркера в конкретному абзаци конкретного твору, з вказівкою на період, жанр і субкорпус (нарративний або діалоговий). Загальні правила бінарного кодування (незалежність семи категорій, множинне кодування як норма — у середньому 1,5 активних категорій на абзац, незалежність від носія ціннісного судження) розгорнуто у Розділі 3 Додатку II. Це відповідає вимозі процедурної прозорості, сформульованій у працях Криппендорфа (Krippendorff, 2019) і Гріммера та інших (Grimmer et al., 2022) як обов'язковій умові валідного словниково-орієнтованого дослідження.

Категоріальна система дослідження.

Категоріальна система побудована на дедуктивно-індуктивному принципі. Дедуктивну рамку становить теорія базових цінностей Шварца (Schwartz, 1992; Schwartz et al., 2012), яку детально розкрито в підрозділі 1.1. У межах цієї теорії десять базових цінностей утворюють мотиваційне коло, структуроване двома ортогональними вищими вимірами: «відкритість до змін» проти «збереження» і «самотрансцендентність» проти «самоствердження». У кодобуку дослідження використано саме цей верхній, агрегований рівень: чотири категорії групи А відповідають чотирьом квадрантам мотиваційного кола.

Індуктивну надбудову становлять три категорії групи В, що відображають специфіку українського соціокультурного контексту 1991–2025 рр. і не зводяться до жодної з універсальних шварцівських категорій. Теоретичним фундаментом групи В є концепції культурної травми (Alexander, 2012; Sztompka, 2000), постсоціалістичної реконструкції національної ідентичності (Offe, 1991) та колективної суб'єктності, актуалізованої під час Революції на граніті 1990 р., Помаранчевої революції 2004 р., Революції Гідності 2013–2014 рр. та повномасштабної війни з 2022 р. (детальніше — у підрозділі 1.2). При цьому категорії групи В не є «цінностями» у вузькому психометричному сенсі, тотожному базовим цінностям Шварца. Вони функціонують як контекстно-ціннісні категорії українського художнього дискурсу: UA-IDN фіксує ціннісну рамку ідентичності, UA-WAR — травматично-пам'яттєву рамку, а UA-FRE — рамку свободи, спротиву й колективної суб'єктності. Зведений склад категорій наведений у Таблиці 2.6.

Таблиця 2.6.

Категоріальна система дослідження

Гр.	Код	Назва	Що агрегує / походження
A	OPN	Відкритість до змін	Самостійність + Стимуляція + Гедонізм (Шварц)
A	CON	Збереження	Безпека + Конформність + Традиція (Шварц)
A	TRA	Самотрансцендентність	Доброзичливість + Універсалізм (Шварц)
A	ENH	Самоствердження	Досягнення + Влада (Шварц)
B	UA-IDN	Національна ідентичність і культурна самобутність	Українська самоідентифікація + культурна пам'ять
B	UA-WAR	Травма війни, переміщення і пам'ять про катастрофи	Військова травма + історичні катастрофи + колективна пам'ять
B	UA-FRE	Свобода і колективна суб'єктність	Національна свобода + спротив + колективна дія

Примітка. Важливо підкреслити, що код TRA у межах цього дослідження не означає Tradition. Традиція як базова цінність III. Шварца входить до агрегованої категорії CON (Conservation / Збереження). Натомість TRA використовується як технічне скорочення для категорії Self-Transcendence / Самотрансцендентність, що агрегує Benevolence і Universalism. Таке позначення зберігається з міркувань стабільності назв змінних у файлах автоматизованого кодування.

Аргументація вибору агрегованих категорій. Вибір чотирьох агрегованих категорій замість десяти базових цінностей Шварца зумовлений потребами словникового аналізу великого корпусу: агрегація підвищує recall і забезпечує стабільніші частоти для міжперіодного порівняння (збільшення обсягу seed-лем у межах однієї категорії підвищує ймовірність виявлення релевантних маркерів); водночас вона зменшує шум між суміжними цінностями, які у самій теорії Шварца функціонально співмірні і часто не диференціюються емпірично навіть на рівні шкальних відповідей (про труднощі емпіричного розрізнення суміжних базових цінностей Шварц пише сам (Schwartz, 1992); ця ж властивість документована у методичних доглядах валідаційних досліджень PVQ-RR (Schwartz et al., 2012)). Для художнього дискурсу тонша диференціація між суміжними цінностями (наприклад, Самостійність окремо від Стимуляції в межах квадранта OPN) погано операціоналізується через лексичні маркери і дає істотний шум; утрата тонкої диференціації компенсується якісним аналізом відібраних кейсів (підрозділ 2.2).

Аргументація окремої групи В. Категорії UA-IDN, UA-WAR і UA-FRE введено не як «українську специфіку» у межах універсальної рамки Шварца, а як самостійні ціннісні утворення з власним теоретичним фундаментом. UA-IDN охоплює ціннісні орієнтири, пов'язані з самовизначенням як українця та з культурною пам'яттю; UA-WAR — досвід війни, репресій, Голодомору, Чорнобиля і переміщення як основу колективного «ми»; UA-FRE — національну (а не індивідуальну) свободу, спротив і революційний вибір. Ці утворення мають колективний, а не індивідуально-мотиваційний характер і тому не вписуються в шварцівське мотиваційне коло. Кодбук передбачає, що категорії груп А і В часто

з'являються разом: мотив самопожертви на війні кодується як TRA + UA-WAR, мотив відмови від російської мови на користь української в умовах війни — як UA-IDN + UA-FRE + UA-WAR.

Розмежування «свободи» між OPN і UA-FRE. Слово «свобода» і його синоніми (вільність, незалежність, воля) є найбільш проблемним випадком кодбуку, оскільки лексично воно належить одночасно до OPN (як індивідуальна автономія) і до UA-FRE (як національна, колективна свобода). Правило розмежування реалізовано не на лексичному рівні, а через автоматично обчислюване поле `subject_level`: за наявності в абзаці маркерів індивідуального суб'єкта («я», «він», «вона», «себе») лема відносить кейс до OPN; за наявності маркерів колективного суб'єкта («ми», «народ», «нація», «українець», «спільнота») — до UA-FRE; у граничних випадках, коли свобода особи постає в умовах національної несвободи, обидві категорії присвоюються одночасно. Повний перелік маркерів суб'єктного рівня, правила обробки граничних значень `subject_level` (mixed, unclear, multivalent), а також каталог типових випадків — у Додатку И (розділ 4.1 і розділ 7).

Повний перелік маркерів, граничних випадків, контекстних правил і процедур розмежування — у Додатку И (кодбук дослідження). У основному тексті обмежуємось архітектурним описом сітки і логікою її побудови.

Технологічна реалізація автоматизованого кодування.

Автоматизована процедура кодування реалізована на стандартному інструментарії обробки українськомовних текстів, доповненому власними скриптами для розширення словника, прогону по корпусу та обчислення метрик. Технологічний стек складається з чотирьох компонентів — лематизації, обчислення тональності, розширення словника й автоматизованих полів

кодування. Загальну архітектуру аналітичної послідовності роботи з текстом та похідними від нього наведено на Рисунку 2.1.

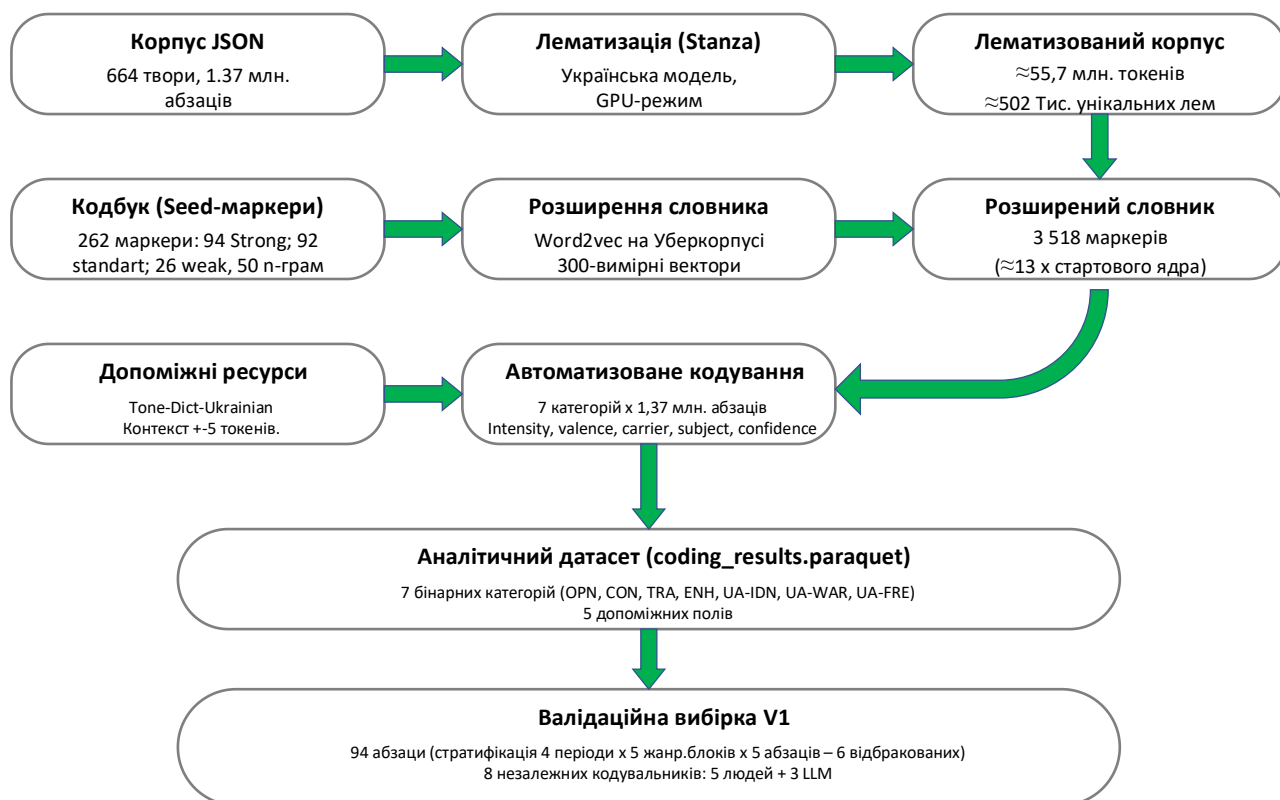


Рис. 2.1. Архітектура аналітичної послідовності автоматизованого аксіологічного контент-аналізу

Лематизація. Корпус і всі seed-маркери кодбуку лематизуються через бібліотеку Stanza (Stanford NLP Group) (Qi et al., 2020) з українською моделлю, навченою на даних Universal Dependencies. Лематизація необхідна для приведення словоформ до базової форми, без чого словник маркерів довелось би розгортати на всі граматичні форми кожної лема. Окрім лем, Stanza-пайплайн зберігає для кожного слова частину мови (UPOS) і морфологічні ознаки, що використовуються при фільтрації кандидатів-маркерів і у якісному аналізі. Для лем, на яких Stanza систематично виводить помилковий результат через омонімію або плюральні форми, передбачено механізм `lemma_overrides` — буквальный пошук словоформи замість лематизованого пошуку (деталі

реалізації, перелік перевизначених лем і умови GPU-прогону — у Додатку И, розділ 6.2).

Тональність. Як допоміжний показник застосовується словник Tone-Dict-Ukrainian (lang-uk, n.d.-a) (3 442 слова з оцінкою емоційної полярності за шкалою $-2...+2$), розроблений командою lang-uk. Tone-Dict використовується для обчислення поля valence — середньої полярності слів у вікні ± 5 токенів від маркера. Принципово важливе застереження: valence не є оцінкою ціннісного знаку, а є показником емоційної полярності лексичного оточення, який корелює з ціннісним знаком, але йому не дорівнює. У роботі valence не використовується як самостійна підстава для присвоєння або зняття категорії, а слугує описовим параметром для якісного аналізу.

Розширення словника. Стартовий словник маркерів (загалом 262 одиниці: strong- і standard-леми, weak-леми та n-грами) було розширено за допомогою предтренуваної word2vec-моделі (Mikolov et al., 2013), навченої на лематизованому Уберкорпусі української мови (Chaplynskyi, 2021; lang-uk, n.d.-b). Для strong- і standard-лем добиралися семантично близькі кандидати, які проходили автоматичну фільтрацію за чотирма показниками: подібністю до seed-леми, нормалізованою корпусною частотою, контекстуальною узгодженістю та специфічністю для категорії. Частина кандидатів приймалася автоматично; багатокатегорійні й семантично неоднозначні випадки передавалися на ручну верифікацію або отримували контекстні правила розмежування залежно від кількості категорій походження та значення композитного score. У результаті фінальний словник було розширено до 3 518 маркерів, що становить ≈ 13 -кратне зростання від стартового ядра. Детальну процедуру оцінювання кандидатів — формулу композитного score, пороги прийняття, статуси кандидатів і правила ручної верифікації — наведено в Додатку И, розділ 6.1. Підсумковий розподіл 4 384 кандидатів за чотирма

статусами автоматичної класифікації (auto_accept, auto_accept_multivalent, manual_review, rejected) подано у Таблиці И.6.1.

Автоматизовані поля кодування. Окрім самого факту присвоєння категорії абзацу, для кожного спрацювання обчислюється набір службових полів: intensity_weighted — зважена сума ваг знайдених маркерів у межах абзацу (ваги: strong-n-грама = 2,0; strong-лема = 1,5; standard-лема = 1,0; weak-лема = 0,5; кількісний показник щільності категорії); valence — за Tone-Dict-Ukrainian; carrier_auto — тип голосу, у якому виражається оцінка; subject_level_auto — рівень суб'єкта висловлювання; confidence — інтегральний діагностичний показник у діапазоні [0; 1]. У межах цієї роботи confidence має допоміжний діагностичний статус і не використовується як самостійна підстава для основних висновків. Деталізовані правила автоматичного присвоєння значень subject_level_auto і пов'язані з ними правила розмежування OPN/UA-FRE наведено в Додатку И, розділ 6.4.

Готовність розширеного словника до повного циклу автоматичного кодування підтверджено попередньою діагностикою покриття seed-маркерів на лематизованому корпусі: 99,0% однослівних seed-лем кодбуку мають принаймні одне входження у корпусі, а категорійне покриття активних лем коливається від 84% (CON) до 93% (ENH); деталі діагностики наведено в Додатку И, розділ 6.5. Спільно ці чотири компоненти технологічного стеку забезпечують масштабований і відтворюваний автоматизований прогон по всьому корпусу та повертають для кожного абзацу набір категорій, ваг, тональних і контекстуальних характеристик, придатний для подальшої агрегації за періодами, жанрами і субкорпусами.

Процедура валідації.

Валідація автоматизованої процедури реалізована на контрольному корпусі V1 у двох комплементарних режимах: оцінка внутрішньої надійності за коефіцієнтом α Криппендорфа (узгодженість незалежних кодувальників) і оцінка зовнішньої валідності через зіставлення з трьома референтними розмітками — Human-gold, LLM-gold та Expanded-gold. Така схема відповідає сучасним рекомендаціям щодо валідації словникових методів у text-as-data-дослідженнях (Grimmer et al., 2022; Krippendorff, 2019; Wilkerson & Casas, 2017).

V1-вибірка є стратифіковано-випадковою вибіркою зі 100 абзаців за схемою 4 періоди $\times$ 5 жанрових блоків $\times$ 5 абзаців на комірку; кожен абзац суцільно кодується за всіма сімома категоріями (бінарне рішення 1/0). V1 призначена для оцінки агрегованої якості процедури в цілому, а не для статистично стабільних оцінок за окремими категоріями; категорійні показники інтерпретуються як діагностичні. Для розрахунку precision, recall і F1 після технічного зіставлення автоматичного й референтного кодування використовується валідна підвибірка з 94 абзаців; причини вилучення шести випадків докладно описано в підрозділі 3.2.

До валідації залучено п'ятьох кодувальників-волонтерів, які пройшли інструктивне навчання на основі кодбуку (Додаток И) і працювали в режимі blind. Детальна процедура ручного кодування (інструмент розмітки, інтерфейс, покрокова процедура, формат експорту CSV, організаційні умови, правила формування «золотого стандарту») описана у Розділі 5 Додатку И і доповнена у Додатку К. Паралельно вибірку розмічено трьома мовними моделями (Claude Opus 4.7, Gemini 2.5 Pro, GPT-4o) за стандартизованим промптом (Додаток К). Вибір α Криппендорфа замість κ Коена зумовлений придатністю α для номінальних даних і кількох незалежних кодувальників (Hayes & Krippendorff, 2007; Krippendorff, 2019).

Референтні розмітки формуються як три gold standards: Human-gold — мажоритарне голосування п'яти людей із порогом $\geq 3/5$; LLM-gold — мажоритарне голосування трьох моделей із порогом $\geq 2/3$; Expanded-gold — об'єднаний консенсус усіх восьми кодувальників із порогом $\geq 4/8$. Включення LLM-консенсусу як паралельної референтної розмітки (Gilardi et al., 2023; Ziems et al., 2024) дозволяє розрізнити контекстну компоненту розбіжностей (автомат проти людської семантичної інтерпретації) і лексичну компоненту (узгодження автомата з LLM на рівні текстових маркерів).

На етапі планування задано орієнтовні пороги якості: precision $\geq 0,70$, recall $\geq 0,60$, F1 $\geq 0,65$ для повністю автоматичного класифікаційного інструменту; для міжкодувальної надійності латентного соціокультурного матеріалу прийнятним вважається α Криппендорфа $\geq 0,67$, бажаним — $\geq 0,80$ (Hayes & Krippendorff, 2007). Ці пороги мають діагностичний статус: у разі їх недосягнення інструмент інтерпретується не як повноцінний семантичний класифікатор цінностей, а як індикаторний словниковий механізм виявлення текстових маркерів, результати якого потребують агрегування, жанрового контролю та якісної інтерпретації.

Окремий контрольний корпус V2 для ітеративного налагодження словника не реалізовувався з огляду на обсяг і часові обмеження магістерського проєкту; його запровадження визначено як рекомендацію до подальшого розвитку методу. Повні результати валідації — попарні α , F1 проти кожного gold standard і синтез висновків — подано в Розділі 3.

Межі автоматизованого підходу.

Експліцитна фіксація меж методу є частиною методологічної прозорості і захищає роботу від інтерпретативних надбудов, не підкріплених застосованим інструментом. Обмеження автоматизованого аксіологічного контент-аналізу, що застосовується у цій роботі, групуються у чотирьох площинах.

Межа 1. Маркери цінностей і межі семантичної інтерпретації.

Дослідження фіксує не цінності читачів, авторів або українського суспільства, а ціннісно марковані одиниці художнього тексту. Автоматизоване кодування виявляє лексичні й n-грамні маркери, але не завжди визначає їхню конкретну семантичну функцію: той самий маркер може бути вжитий іронічно, заперечно, метафорично, у мовленні персонажа або в інтертекстуальному посиленні. Перехід від маркерів до соціологічної інтерпретації опосередкований авторським задумом, жанровими конвенціями та редакторсько-видавничими фільтрами. Тому результати автоматизованого кодування не інтерпретуються на рівні окремого абзацу як остаточне смислове рішення; їхня основна функція — побудова агрегованих індикаторів, порівнюваних між періодами, жанровими блоками й субкорпусами. Зокрема, контроль домінування G17 (історична проза, 21,2% корпусу — див. підрозділ 2.1) необхідний для уникнення приписування жанрової специфіки загальній динаміці ціннісних маркерів.

Межа 2. Маніфестний і латентний зміст. Автоматизована процедура виявляє переважно маніфестний (експліцитно виражений) ціннісний зміст — оцінні слова, прямі моральні судження, лексично марковані висловлювання персонажів. Латентний шар може бути частково врахований через якісну інтерпретацію показових кейсів, відібраних за результатами автоматизованого кодування; у межах цієї роботи такий аналіз має допоміжний, ілюстративно-інтерпретативний характер. Для цього аналізу використовуються наративні й символічні маркери довідково-інтерпретативного блоку кодбуку (розділ 7 Додатку II), які не застосовуються у масовому автоматизованому кодуванні.

Межа 3. Межі ручної та LLM-валідації. Зовнішня валідація автоматизованого кодування спирається на дві опори — ручну розмітку п'ятьох кодувальників-інженерів і LLM-розмітку трьох мовних моделей. Обидві мають характеристики, що накладають обмеження на висновки.

На стороні ручної розмітки: п'ять кодувальників забезпечують робочу схему для мажоритарного голосування і розрахунку α Криппендорфа, однак такий масштаб не забезпечує статистичної стабільності окремих категорійних метрик; за низькочастотними категоріями кількість позитивних кейсів у вибірці V1 недостатня для надійних оцінок. Окремою обставиною є те, що в інструктивному навчанні кодувальникам подавалися функціональні (операційні) визначення категорій, а не їх повна теоретична розгортка з опорою на праці Шварца, Дюркгейма, Бурдьє та ін. Така процедурна організація узгоджується зі стандартами навчання волонтерських кодувальників (Krippendorff, 2019; Lombard et al., 2002; Neuendorf, 2017), але водночас означає, що тонкі розрізнення між суміжними категоріями (зокрема індивідуального OPN і колективного UA-FRE, побутового CON і національного UA-IDN) ґрунтуються переважно на лексично-функціональному рівні розпізнавання, що частково пояснює низькі α -показники людської узгодженості саме для цих категорій.

На стороні LLM-розмітки: три моделі демонструють вищу узгодженість між собою на лексико-семантичному рівні, але це не гарантує глибшого семантичного розрізнення в неоднозначних художніх контекстах. Усі висновки роботи формуються з відповідним епістемічним застереженням: ідеться про тенденції, виявлені інструментом заявленої якості, а не про беззаперечні факти ціннісної структури дискурсу.

Межа 4. Іронія, фокалізація, невластне-пряме мовлення та інтертекст. Окремий клас обмежень пов'язаний з художнім модусом текстів. Іронія, сарказм, невластне-пряме мовлення, цитати з інших авторів, метафоричні розширення ключових слів і складна наративна фокалізація — усі ці явища ускладнюють автоматичне кодування. Кодбук містить специфічні правила для типових випадків (каталог граничних випадків — у розділі 8 Додатку II), однак повна інтерпретація іронії і латентної авторської позиції залишається у площині якісного аналізу. Особливо це стосується творів постмодерністської і

експериментальної прози, де іронічна дистанція до ціннісних маркерів є структурним прийомом.

Сукупно зазначені межі окреслюють поле дослідження як аналіз динаміки текстових маркерів ціннісних категорій у конкретному корпусі, а не як пряме вимірювання цінностей суспільства. Отримані кількісні показники набувають соціологічного значення лише в поєднанні з перевіркою валідності, жанровим контролем, аналізом трендів і якісною інтерпретацією показових випадків. Саме тому в Розділах 3–4 результати автоматизованого кодування розглядаються як індикативні, а не як самодостатній доказ трансформації ціннісних орієнтацій.

2.3. Висновки до розділу 2

У Розділі 2 сформовано емпіричну й інструментальну основу дослідження, яка визначає межі подальшої інтерпретації результатів. Підготовлений корпус україномовної художньої прози 1991–2025 рр. є не репрезентативною вибіркою всієї української літератури, а цільовим аналітичним масивом із 664 творів 235 авторів, структурованим за хронологічними зрізами, жанровими блоками та субкорпусами наративу й діалогу. Така структура забезпечує достатній масштаб для корпусного аналізу, але водночас вимагає постійного контролю інституційного, жанрового, часового й технічного зміщення.

Розроблена категоріальна система поєднує агреговані виміри моделі Ш. Шварца з контекстно-специфічними категоріями українського художнього дискурсу. Це дозволяє аналізувати як універсальні ціннісні опозиції, так і специфічні для українського досвіду рамки ідентичності, війни, пам'яті, свободи й колективної суб'єктності. Водночас така категоріальна сітка не є психометричним інструментом вимірювання цінностей, а виступає операційною моделлю фіксації текстових маркерів.

Автоматизований інструмент кодування забезпечує масштабне виявлення ціннісно маркованих одиниць, але його результати мають індикативний характер. Словниковий підхід фіксує насамперед маніфестний лексико-семантичний рівень і не здатний повністю реконструювати іронію, латентну авторську позицію, наративну фокалізацію чи інтертекстуальні смисли. Тому результати автоматизованого кодування потребують перевірки валідності, жанрової стандартизації й якісної інтерпретації показових випадків.

Отже, Розділ 2 задає не лише технічну процедуру аналізу, а й режим допустимих висновків: у подальших розділах ідеться про динаміку текстових маркерів ціннісних репрезентацій у сформованому корпусі, а не про пряме вимірювання цінностей українського суспільства, авторів або читачів.

РОЗДІЛ 3. ВАЛІДАЦІЯ АВТОМАТИЗОВАНОГО АКСІОЛОГІЧНОГО КОДУВАННЯ

Валідаційний блок дослідження побудований на двох комплементарних рівнях, які разом забезпечують методологічну прозорість автоматизованого аксіологічного контент-аналізу. Внутрішня надійність (підрозділ 3.1) оцінює послідовність кодувальних рішень між незалежними учасниками — як людьми, так і LLM-моделями — і вимірюється через коефіцієнт α Криппендорфа. Зовнішня валідність (підрозділ 3.2) оцінює відповідність автоматичного словникового методу трьом незалежним «золотим стандартам», побудованим через мажоритарне голосування. Поєднання двох рівнів дозволяє розрізнити два типи розбіжностей: контекстну (де автомат відхиляється від семантичного розуміння людини) і лексичну (де автомат і LLM узгоджені між собою, бо обидва спираються на текстові маркери) (Gilardi et al., 2023; Ziems et al., 2024).

3.1. Внутрішня надійність кодування: α Криппендорфа

Методика обчислення. Для оцінки внутрішньої надійності V1-вибірка (100 абзаців $\times$ 7 категорій = 700 одиниць) була незалежно закодована кількома учасниками-людьми (A, L, N, V, T) у режимі сліпого кодування. Паралельно та сама вибірка прокодована трьома LLM-моделями (Claude Opus 4.7, Gemini 2.5 Pro, GPT-4o) через автоматичний інтерфейс зі стандартизованим промптом (Додаток К), побудованим на основі функціональної інструкції кодбуку (Додаток И). Для кожної з семи бінарних категорій обчислено α Криппендорфа за класичною формулою для номінальних даних (Hayes & Krippendorff, 2007; Krippendorff, 2019). Інтерпретація шкали α : $\geq 0,80$ — високе узгодження; $0,67 \leq$

$\alpha < 0,80$ — прийнятне; $0,40 \leq \alpha < 0,67$ — слабе; $\alpha < 0,40$ — фактично відсутнє узгодження.

Загальна α серед п'яти кодувальників-людей становить 0,265, що за прийнятою у цій роботі шкалою інтерпретації відповідає фактично відсутньому узгодженню. Серед трьох LLM-моделей загальна α становить 0,588, а середня попарна α LLM–LLM — 0,590. Оскільки обидва показники практично збігаються, у змістовній інтерпретації можна говорити про рівень близько 0,59, але в таблицях слід розрізняти загальну багатокодувальну α і середню попарну α . Розрив між загальною α людей і LLM становить 0,323 пункти α , що вказує на різні режими застосування кодбуку: людські кодувальники активніше враховують контекст, іронію, наративну позицію та підтекст, тоді як LLM-моделі демонструють більшу стабільність у розпізнаванні лексико-семантичних патернів (Gilardi et al., 2023; Hayes & Krippendorff, 2007; Krippendorff, 2019; Ziems et al., 2024). Детальніше дані про α Криппендорфа надані у Таблиці 3.1.

Таблиця 3.1.

α Криппендорфа за сімома ціннісними категоріями: п'ять кодувальників-людей і три LLM-кодувальники

Категорія	α людей	Інтерпретація	α LLM	Інтерпретація	Коментар
UA-IDN	0,512	слабе узгодження	0,568	слабе узгодження	Найкраща категорія серед людей
UA-WAR	0,236	фактично відсутнє	0,505	слабе узгодження	Людська інтерпретація нестабільна
TRA	0,236	фактично відсутнє	0,708	прийнятне узгодження	Найкраща категорія серед LLM
UA-FRE	0,230	фактично відсутнє	0,650	слабе, близьке до прийнятного	Потребує уточнення меж категорії

OPN	0,229	фактично відсутнє	0,532	слабке узгодження	Висока контекстна неоднозначність
ENH	0,181	фактично відсутнє	0,675	прийнятне узгодження	LLM краще вловлюють лексичні маркери
CON	0,181	фактично відсутнє	0,388	фактично відсутнє	Найпроблемніша категорія

Інтерпретація профілю α по категоріях. Порівняння α за окремими категоріями показує, що після включення п'ятого людського кодувальника відносно найкраще розпізнаваною серед людей залишається категорія UA-IDN: $\alpha = 0,512$. Це означає, що маркери національної ідентичності, мови, культурної самобутності та історичної пам'яті є відносно більш упізнаваними для людських кодувальників. Натомість усі інші категорії серед людей мають α нижче 0,40, тобто не демонструють достатньої міжкодувальної стабільності. Особливо проблемними є CON та ENH, де розмежування між збереженням, владою, досягненням, порядком і соціальним контролем часто потребує ширшої контекстної інтерпретації.

Профіль LLM-кодування є іншим. Найвищі значення α спостерігаються для TRA (0,708) та ENH (0,675), що свідчить про кращу машинну стабілізацію категорій, насичених відносно чіткими лексичними маркерами турботи, допомоги, досягнення, сили або статусу. Принципово важливим є те, що для TRA — категорії, яка серед людей має фактично відсутнє узгодження ($\alpha = 0,236$) — LLM-моделі досягають прийнятного рівня узгодження за стандартами Криппендорфа (Hayes & Krippendorff, 2007; Krippendorff, 2019; Lombard et al., 2002). Цей розрив у +0,472 пункти α між групами для однієї і тієї самої категорії емпірично обґрунтовує методичну роль LLM-консенсусу як паралельного gold standard у підрозділі 3.2: там, де людське кодування демонструє низьку міжкодувальну узгодженість, LLM-моделі стабільніше реагують на лексико-семантичний шар тексту. Це не означає вищої змістової валідності LLM, але

показує їхню корисність як окремої діагностичної рамки. Водночас CON залишається проблемною і для LLM-моделей ($\alpha = 0,388$), що підтверджує структурну складність цієї категорії в художньому матеріалі.

Ключова асиметрія: Human–Human проти LLM–LLM. Серед десяти можливих пар кодувальників-людей середня попарна α становить 0,274 (діапазон 0,119–0,438). Серед трьох пар LLM середня попарна α становить 0,590 (діапазон 0,561–0,634). Розрив між середніми попарними показниками становить 0,316 пункти α , тобто LLM-кодування демонструє істотно вищу формальну стабільність у межах заданого кодбуку. Жодна з десяти пар «людина–людина» не перевищує мінімального значення серед пар «LLM–LLM» (0,561 - детальніше у Таблиці 3.2). Повну попарну α -матрицю узгодженості між усіма вісьмома кодувальниками — п'ятьма людьми (A, L, N, T, V) і трьома LLM-моделями (Claude Opus 4.7, Gemini 2.5 Pro, GPT-4o) — подано в Таблиці Л.1 Додатка Л. Це методично важливий результат, який обґрунтовує включення LLM-консенсусу як паралельного gold standard у підрозділі 3.2. Водночас цю асиметрію не слід інтерпретувати як «кращу» змістову інтерпретацію з боку LLM: вона радше свідчить про більшу формальну узгодженість моделей у розпізнаванні лексико-семантичних патернів, тоді як людське кодування сильніше залежить від контексту, іронії, наративної перспективи та латентного смислу. Така більша формальна стабільність LLM у задачах текстової анотації узгоджується з результатами сучасних досліджень використання великих мовних моделей як кодувальників (Gilardi et al., 2023; Ziems et al., 2024).

Таблиця 3.2.

Середні попарні α у внутрішньо однорідних групах

Тип пар	Кількість пар	Середня α	Діапазон	SD
Human--Human	10	0,274	0,119–0,438	0,119

LLM--LLM	3	0,590	0,561–0,634	0,039
Human--LLM	15	0,327	0,170–0,458	0,093

Примітка. Діагностична непараметрична перевірка різниці між парами Human--Human і LLM--LLM показує статистично виражений розрив: $U = 0,00$; $p = 0,0035$. Водночас цей результат слід трактувати обережно, оскільки попарні коефіцієнти не є повністю незалежними спостереженнями: ті самі кодувальники входять у кілька пар.

Інтерпретація низької α людей як методичного обмеження. Низький рівень α серед кодувальників-людей (0,265) не може трактуватися як достатній рівень міжкодувальної надійності у строгому сенсі. Він свідчить про високу складність операціоналізації латентного аксіологічного змісту художнього тексту, де рішення залежать від контексту, іронії, метафори, наративної перспективи та межі між маніфестним і латентним значенням. Тому результати людського кодування в цьому дослідженні використовуються не як безумовний еталон, а як один із трьох діагностичних стандартів поряд із LLM-gold та Expanded-gold. Така позиція відповідає сучасним рекомендаціям щодо валідації автоматизованих процедур в умовах кодування латентного смислу: за межами маніфестних формальних ознак (тематика, поверхневі лексичні маркери) орієнтація лише на людську α -узгодженість без паралельних діагностичних стандартів призводить до недооцінки інструменту або, навпаки, до невиправдано оптимістичних висновків (Lombard et al., 2002; Neuendorf, 2017; Wilkerson & Casas, 2017). Саме поєднання кількох стандартів дозволяє не усунути, а явно зафіксувати різницю між контекстно-інтерпретативною і лексико-семантичною логікою кодування (Hayes & Krippendorff, 2007; Krippendorff, 2019; Lombard et al., 2002).

При мажоритарному порозі ≥ 3 з 5 кодувальників випадки з найслабшою консенсусністю (рівно 3 голоси з 5, тобто мінімальна кваліфікаційна більшість) фіксуються у підсумкових таблицях окремо — вони мають вищу невизначеність,

ніж випадки однастайності 4–5 голосів. У підсумкових метриках Розділу 3 ці випадки включаються в $\text{gold}=1$ без окремої adjudication-процедури.

3.2. Зовнішня валідність: F1 проти трьох gold standards

Оцінка узгодженості словникового методу. Зовнішня валідність автоматичного словникового методу оцінюється через precision, recall і F1 проти трьох незалежних gold standards, побудованих через мажоритарне голосування (Grimmer et al., 2022; Neuendorf, 2017). Розрахунок здійснювався не на всій V1-вибірці, а на валідній підвибірці з 94 абзаців: із 100 початкових позицій 6 було виключено з технічних причин, зокрема три випадки мали колізії ідентифікаторів абзаців, а ще три позиції не були зіставлені з автоматичним прогоном. Таким чином, зовнішня валідизація охоплює 658 бінарних рішень: $94 \text{ абзаци} \times 7 \text{ категорій}$.

Три gold standards формують ієрархію оцінювальних рамок: **Human-gold** — мажоритарне голосування 5 кодувальників-людей (A, L, N, T, V) з порогом ≥ 3 з 5: **F1 = 0,076**; **LLM-gold** — мажоритарне голосування 3 LLM-кодувальників (Claude Opus 4.7, Gemini 2.5 Pro, GPT-4o) з порогом ≥ 2 з 3: **F1 = 0,202**; **Expanded-gold** — об'єднаний консенсус 8 кодувальників (5 людей + 3 LLM) з порогом ≥ 4 з 8: **F1 = 0,119**.

Отримані значення F1 є низькими для всіх трьох gold standards і не дозволяють трактувати словниковий алгоритм як надійний семантичний класифікатор окремих абзаців. Найвищий результат спостерігається щодо LLM-gold (F1 = 0,202), тоді як відповідність Human-gold становить лише F1 = 0,076. Це означає, що автоматичний метод значно краще узгоджується з лексико-семантично орієнтованим машинним консенсусом, ніж із контекстно-інтерпретативним людським кодуванням. Отже, його слід використовувати не як

інструмент точної мікрокласифікації, а як макрорівневий індикатор активності ціннісних маркерів у великому корпусі.

У подальшому викладі слід розрізняти дві бази розрахунку. Повна V1-вибірка містить $100 \text{ абзаців} \times 7 \text{ категорій} = 700$ бінарних рішень і використовується для опису загальної активності трьох gold standards. Натомість метрики TP, FP, FN, TN, precision, recall і F1 обчислено на валідній підвибірці з 94 абзаців, тобто на 658 бінарних рішеннях. Саме цим пояснюється різниця між кількістю позитивних випадків у таблиці 3.3 і відсотками активності в таблиці 3.4.

Узгодженість словникового методу з LLM-консенсусом у **2,66 рази** вища за узгодженість з людським консенсусом. При цьому Human-gold виявився надзвичайно консервативним: на повній V1-вибірці (700 одиниць) лише 35 (5%) отримали мажоритарну мітку «1»; на валідній підвибірці (658 одиниць) — 34 (5,2%). Автоматичний метод активувався в 24% одиниць у середньому. Особливо мало позитивних випадків у людському консенсусі для категорій UA-WAR (2 з 100) та UA-FRE (1 з 100), що робить статистичну оцінку F1 у цих категоріях вкрай чутливою до випадкових варіацій. LLM-gold виявився у 3 рази щедрішим (15%) і ближчим до автоматичного методу за лексичними патернами активації, що відповідає висновкам (Gilardi et al., 2023; Ziems et al., 2024) про функціональну подібність LLM-консенсусу і автоматичних методів у задачах кодування латентного семантичного змісту.

Серед семи ціннісних категорій найкращу узгодженість показала OPN ($F1=0,369$ vs LLM-gold; $F1=0,208$ vs Human-gold), найгіршу — UA-IDN ($F1=0$ проти всіх трьох gold standards). Повний розклад метрик TP, FP, FN, TN, precision, recall і F1 за кожною з семи ціннісних категорій і кожним із трьох gold standards (всього 21 рядок) наведено в Таблиці М.1 Додатка М; саме звідти отримано підсумкові показники, наведені в Таблицях 3.3 і 3.5 основного тексту. Особлива природа цього розриву зумовлена структурною неузгодженістю двох

рівнів валідації: за α -показником людської узгодженості UA-IDN є найкращою категорією серед людей ($\alpha = 0,512$, таблиця 3.1), проте автоматичне кодування активується на лексичних патернах, які не збігаються з людським, LLM- або Expanded-консенсусом. Це підтверджує необхідність структурної ревізії словника UA-IDN у напрямку сучасного шару української ідентичнісної лексики 2014–2025 рр., що узгоджується з обмеженнями автоматизованого підходу, сформульованими у підрозділі 2.2.

Реальні значення F1 нижчі за початково цільові пороги, прийняті у підрозділі 2.2 ($\text{precision} \geq 0,70$; $\text{recall} \geq 0,60$; $F1 \geq 0,65$), що відображає відому асиметрію між мікрорівнем (поодинокий абзац) і макрорівнем (агрегатні корпусні показники) словникового методу у задачах кодування художнього дискурсу (Krippendorff, 2019; Wilkerson & Casas, 2017). Інтерпретація результатів у наступних розділах ґрунтується на агрегатних показниках, для яких такий рівень одиничної F1 не є дискваліфікуючим, якщо систематичні зміщення не корелюють з часом і жанром.

Таблиця 3.3 фіксує ключову асиметрію зовнішньої валідності: автоматичний метод краще узгоджується з LLM-gold, ніж із Human-gold.

Таблиця 3.3.

Підсумкові метрики зовнішньої валідності автоматичного кодування проти трьох gold standards у валідній підвибірці (94 абзаци $\times$ 7 категорій = 658 рішень)

Gold standard	Gold=1 у валідній підвибірці, 94 $\times$ 7	TP	FP	FN	TN	Precision	Recall	F1
Human-gold	34	7	144	27	480	0,046	0,206	0,076
LLM-gold	96	25	126	71	436	0,166	0,26	0,202
Expanded-gold	68	13	138	55	452	0,086	0,191	0,119

Таблиця 3.4 пояснює одну з ключових причин низьких F1-метрик: Human-gold виявився надзвичайно консервативним — лише 35 позитивних випадків з 700 (5%), причому для категорій UA-WAR і UA-FRE — відповідно 2 і 1 позитиви на 100 абзаців. При такій рідкості позитивів статистична оцінка F1 у цих категоріях стає вкрай чутливою до випадкових варіацій. LLM-gold у 3 рази частіше активує позитивні мітки, ніж Human-gold (15% проти 5%), і тому за загальним рівнем активності ближчий до автоматичного словникового методу. Проте це не означає, що LLM-gold є “точнішим”: він репрезентує іншу логіку кодування, більш чутливу до лексико-семантичних патернів і менш консервативну щодо латентного контексту.

Таблиця 3.4.

Активність ціннісних категорій у трьох gold standards у повній V1-вибірці (100 абзаців × 7 категорій = 700 рішень), % позитивних випадків

Gold	OPN	CON	TRA	ENH	UA-IDN	UA-WAR	UA-FRE	Усього 1's
Human	11.0%	3.0%	5.0%	9.0%	4.0%	2.0%	1.0%	5.0%
LLM	30.0%	11.0%	13.0%	24.0%	7.0%	14.0%	6.0%	15.0%
Expanded	21.0%	8.0%	10.0%	20.0%	8.0%	4.0%	1.0%	10.3%

Таблиця 3.5 демонструє системну характеристику словникового методу — переоцінку присутності ціннісних категорій у 3–10 разів порівняно з людським консенсусом. Найбільший розрив — для категорій UA-FRE (10×), UA-WAR (9×) і CON (8×). Це не лише недолік конкретної реалізації, а й структурна особливість лексично орієнтованого підходу: автомат активується на наявності маркерів і не завжди враховує контекстну семантику, тоді як кодувальник інтегрує іронію, риторичні інверсії, позицію мовця та символічні підтексти. Водночас масштаб розриву за окремими категоріями, особливо UA-IDN, UA-FRE і CON, вказує також на необхідність подальшої ревізії словника.

Таблиця 3.5.

Систематична надмірність автоматичного кодування відносно gold standards у валідній підвибірці (n = 94 абзаци)

Категорія	Auto %	Human-gold %	LLM-gold %	Auto/Human	Auto/LLM
OPN	39.4	11.7	29.8	3.4×	1.3×
CON	25.5	3.2	10.6	8.0×	2.4×
TRA	23.4	5.3	12.8	4.4×	1.8×
ENH	28.7	8.5	23.4	3.4×	1.2×
UA-IDN	13.8	4.3	6.4	3.2×	2.2×
UA-WAR	19.1	2.1	12.8	9.0×	1.5×
UA-FRE	10.6	1.1	6.4	10.0×	1.7×

Найкращу узгодженість з усіма трьома gold standards показала OPN (F1=0,369 vs LLM-gold) — лексично «прозора» категорія з відносно стабільним маркерним набором. Найгірший результат зафіксовано для UA-IDN: у всіх трьох порівняннях TP=0, тому F1 не визначається і фактично свідчить про відсутність істинно позитивних збігів автоматичного кодування з gold standards. Це критичний діагностичний сигнал: автомат активує UA-IDN на наборі лексичних патернів, які не підтверджуються ні людьми, ні LLM, що вказує на застарілість словника UA-IDN та необхідність його ревізії під сучасний шар української ідентичної лексики 2014–2025 рр.

Таблиця попарного співпадіння трьох gold standards (Додаток Н) має лише допоміжний діагностичний характер. Через різку нерівномірність класів overall agreement значною мірою формується збігами за негативним класом (0), тому високі значення загального збігу не суперечать низьким F1-метрикам. Отже, цей показник не використовується як основний критерій валідності; пріоритетними залишаються precision, recall і F1.

3.3. Висновки до розділу 3

У Розділі 3 здійснено двоступеневу валідацію автоматизованого аксіологічного кодування — оцінку внутрішньої надійності за коефіцієнтом α Криппендорфа та зовнішньої валідності за метрикою F1 проти трьох незалежних gold standards, — що дозволило визначити межі довіри до результатів автоматичного словникового методу в наступному розділі.

На рівні внутрішньої надійності виявлено різку асиметрію між людським і LLM-кодуванням. Загальна α п'яти кодувальників-людей становить 0,265 (середня попарна α Human–Human — 0,274), тоді як для трьох LLM-моделей загальна α — 0,588 (середня попарна α LLM–LLM — 0,590). Цей розрив у 0,316 пункти між попарними коефіцієнтами свідчить про два режими застосування кодбуку: людські кодувальники активніше враховують контекст, іронію та наративну позицію, тоді як LLM-моделі стабільніше реагують на лексико-семантичні патерни. Категорія UA-IDN є відносно найкраще розпізнаваною людьми ($\alpha = 0,512$), а TRA та ENH — LLM-моделями ($\alpha = 0,708$ і $0,675$ відповідно); CON залишається проблемною на обох рівнях.

На рівні зовнішньої валідності F1 автоматичного методу становить 0,076 проти Human-gold, 0,202 проти LLM-gold і 0,119 проти Expanded-gold. Узгодженість автомата з LLM-консенсусом у 2,66 раза вища за узгодженість із людським консенсусом, що відтворює виявлену на рівні α асиметрію між двома логіками кодування. Автоматичний метод систематично переоцінює присутність ціннісних категорій у 3–10 разів порівняно з людським консенсусом; найбільший розрив — для UA-FRE (10×), UA-WAR (9×) і CON (8×). Найкращу узгодженість продемонструвала категорія OPN (F1 = 0,369 vs LLM-gold), найгіршу — UA-IDN з TP = 0 в усіх трьох порівняннях. Цей результат має не лише технічне, а й самостійне соціологічне значення: він емпірично свідчить про те, що

автоматизований словник UA-IDN, побудований переважно на класичному ідентифікаційному шарі (Шевченко, козак, тризуб, вишиванка, національні символи), не покриває сучасні форми української ідентифікації, які після 2014 і 2022 рр. реалізуються через практики спротиву, громадянської суб'єктності, мовну переорієнтацію, мовчазну нормалізацію українськості та сюжетно-наративні структури. Структурна ревізія словника UA-IDN, таким чином, є не лише методичним покращенням, а й операційною відповіддю на зафіксовану зміну режиму репрезентації національної ідентичності в художньому дискурсі.

Синтез двох рівнів валідації обґрунтовує діагностичну, а не повністю класифікаційну валідність розробленого інструментарію. Автоматичне кодування не може використовуватися як прямий еквівалент ручної інтерпретації окремих абзаців; його результати інтерпретуються виключно на агрегатному корпусному рівні як індикатори макрорівневої активності ціннісних маркерів. Емпірично підтверджено доцільність гібридної валідаційної стратегії (Human + LLM + Expanded gold): жоден із трьох стандартів окремо не дає цілісної картини, оскільки Human-gold надмірно консервативний для рідкісних категорій, LLM-gold функціонально подібний до автомата за лексико-семантичною логікою активації, а Expanded-gold займає проміжне положення.

Як наслідок, подальший аналіз у Розділі 4 потребує трьох запобіжників — зіставлення сирих трендів із жанровою структурою корпусу, жанрової стандартизації для відмежування ціннісної динаміки від жанрового артефакту та якісної перевірки показових і контрастних кейсів. Без цих процедур макрорівневі тренди можуть бути помилково інтерпретовані як зміна ціннісних орієнтацій, тоді як частково вони можуть відображати зміну жанрової композиції корпусу або структурні особливості автоматичного словникового підходу.

РОЗДІЛ 4. ТРАНСФОРМАЦІЯ ЦІННІСНИХ ОРІЄНТАЦІЙ У ХУДОЖНЬОМУ ДИСКУРСІ 1991– 2025 РР.

У цьому розділі результати автоматизованого аксіологічного кодування корпусу, описаного в Розділі 2 і перевіреного через процедури валідації Розділу 3, розглядаються як емпірична основа для аналізу динаміки ціннісних репрезентацій у дискурсі україномовної художньої прози 1991–2025 рр. Логіка розділу є поетапною: первинний аналіз сирих агрегованих показників активності семи ціннісних категорій >> жанрова стандартизація >> перевірка теоретичних очікувань Н1–Н4 (підрозділ 1.5) >> якісна інтерпретація показових і контрастних кейсів >> діагностичне зіставлення з WVS/EVS і національними опитуваннями >> узагальнення результатів та фіксація обмежень їх інтерпретації.

Усі подальші показники мають операціональний характер: вони описують активність маркерів, співвіднесених із сімома категоріями кодбука, а не «наявність цінностей» як таку. З огляду на результати валідації Розділу 3, відсотки нижче інтерпретуються як порівняльні індикатори активності маркерів, а не як точні оцінки поширеності відповідних цінностей. TRA у цьому розділі послідовно означає Self-Transcendence / самотрансцендентність (доброзичливість + універсалізм), а не Tradition; це принципово для інтерпретації НЗ.

4.1. Первинна динаміка ціннісних категорій у корпусних даних

Підрозділ 4.1 подає первинний опис результатів автоматизованого кодування без корекції на жанрову композицію корпусу. Його завдання — показати картину, яку дає пряме агрегування абзаців у межах кожного хронологічного зрізу. Оскільки корпус має нерівномірну жанрову структуру, ці результати розглядаються як діагностичні й потребують жанрового контролю, здійсненого у підрозділі 4.2.

Розмір і структура корпусу за хронологічними зрізами.

Аналітичний корпус, використаний у розрахунках цього розділу, містить 664 твори 235 авторів, представлені на рівні 1 375 333 абзаців. Розподіл матеріалу між хронологічними зрізами є нерівномірним, але достатнім для макрорівневих порівнянь: 1991–2003 рр. охоплюють 135 творів і 234 972 абзаци; 2004–2013 рр. — 224 твори і 408 358 абзаців; 2014–2021 рр. — 231 твір і 514 916 абзаців; 2022–2025 рр. — 74 твори і 217 087 абзаців. Найкоротший четвертий зріз потребує особливої обережності, оскільки він охоплює лише чотири роки та залежить від цифрової доступності найновіших текстів.

Одиницею нормалізації в цьому розділі виступає абзац. Для кожної категорії активність розраховувалася як частка абзаців відповідного періоду, у яких автоматизована процедура активувала принаймні один маркер цієї категорії. Така нормалізація дозволяє порівнювати періоди з різним абсолютним обсягом текстового матеріалу, але не усуває впливу жанрового складу, тому результати 4.1 слід розглядати як діагностичні.

Активність семи ціннісних категорій у чотирьох періодах.

Зведена картина сирової активності семи ціннісних категорій подана в таблиці 4.1. Вона показує частку абзаців з активованою категорією в межах кожного хронологічного зрізу без жанрової стандартизації.

Таблиця 4.1.

Активність семи ціннісних категорій у чотирьох хронологічних зрізах корпусу, % абзаців з активованою категорією.

Категорія	1991–2003	2004–2013	2014–2021	2022–2025	Δ (4 vs 1), п.п.
OPN	23,01	21,43	22,20	21,29	–1,72
CON	20,41	16,90	17,71	15,56	–4,85
TRA	15,03	13,62	13,92	13,77	–1,26
ENH	17,12	15,81	16,70	15,83	–1,29
UA-IDN	13,04	8,88	8,70	7,04	–6,00
UA-WAR	14,04	10,96	12,61	11,56	–2,48
UA-FRE	8,36	6,49	6,57	6,11	–2,25

Примітка. Розраховано за результатами автоматизованого кодування coding_results.parquet. Активність означає частку абзаців, у яких категорія отримала значення cat = 1. Через округлення різниці можуть незначно відрізнятися від розрахунків за нерозокругленими значеннями. Джерело t_03_trends_by_period_v2_v2.ipynb

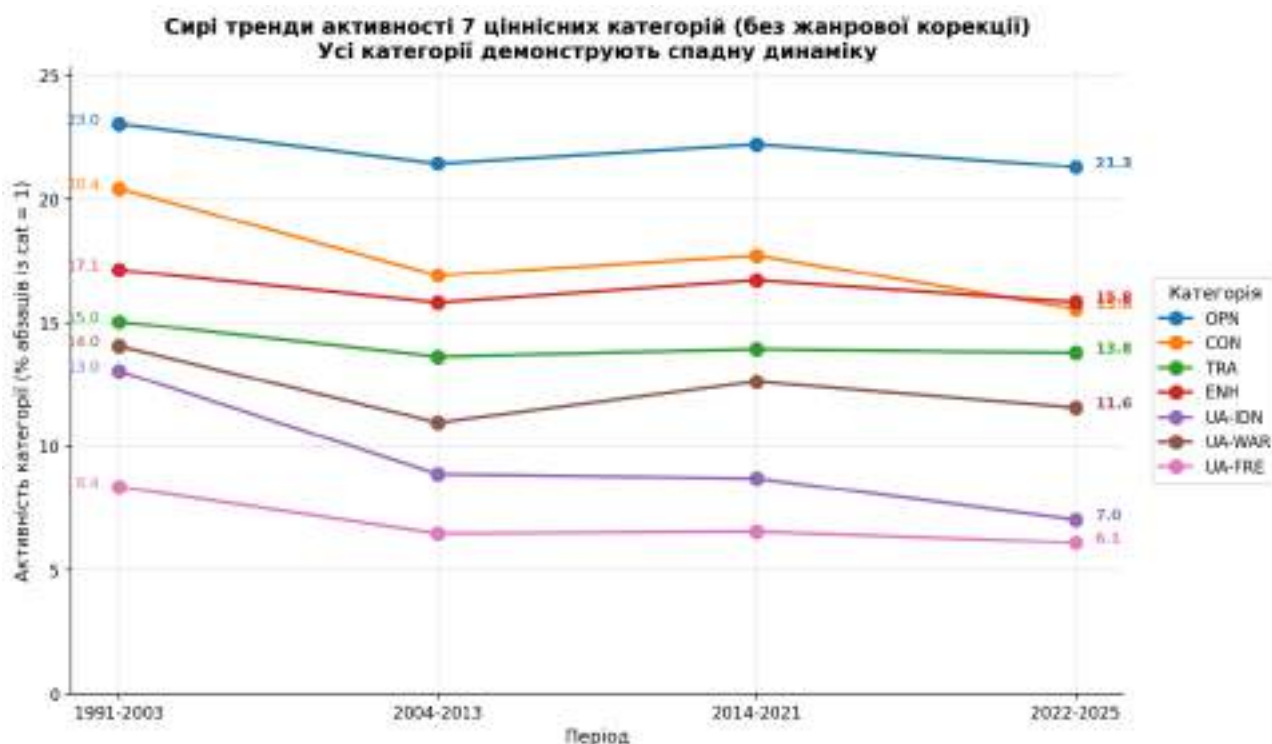


Рис. 4.1. Сирі тренди активності семи ціннісних категорій у чотирьох хронологічних зрізах.

Усі сім категорій демонструють нижчий рівень активності у четвертому періоді (Табл. 4.1, остання колонка). Найбільший спад — для UA-IDN (–6,00 п.п.) і CON (–4,85 п.п.); найменший — для TRA та ENH (–1,26 і –1,29 п.п.).

Первинно така картина могла б бути витлумачена як загальне послаблення ціннісної насиченості художнього дискурсу. Проте саме системність спадного патерну — падіння всіх семи категорій одночасно — вимагає перевірки на наявність структурного артефакту. Малоімовірно, що всі категорії, включно з концептуально протилежними полюсами моделі Шварца (Schwartz, 1992; Schwartz et al., 2012), синхронно знижуються лише внаслідок власне ціннісної трансформації — мотиваційне коло Шварца передбачає, що зростання активності одного полюсу типово супроводжується ослабленням протилежного, а не їх синхронним згасанням.

Перевірка трендів: Cochran–Armitage test та додаткові статистичні показники.

Для перевірки наявності монотонного тренду в упорядкованих хронологічних зрізах застосовано Cochran–Armitage trend test (Armitage, 1955; Cochran, 1954) (результати – дивись Таблицю 4.2). Тест перевіряє нульову гіпотезу про відсутність лінійної залежності між порядком періоду і часткою активних абзаців. Від’ємне значення Z-статистики означає спадну тенденцію, додатне — зростальну. Додатково для загальної асоціації між періодом і бінарною активністю категорії наведено Cramer’s V. Його включено для запобігання надмірній інтерпретації p-value, оскільки за обсягу понад 1,3 млн абзаців навіть дуже малі відмінності стають статистично значущими (Cramér, 1946; Lin et al., 2013).

Таблиця 4.2.

Cochran–Armitage test на наявність монотонного тренду активності ціннісних категорій між чотирма хронологічними зрізами

Категорія	Z-статистика	p-value	Знак тренду	Cramer’s V	Інтерпретація
OPN	–9,02	<0,001	спадний	0,0146	статистично значущий, дуже слабкий
CON	–34,88	<0,001	спадний	0,0386	статистично значущий, слабкий за силою
TRA	–9,79	<0,001	спадний	0,0140	статистично значущий, дуже слабкий
ENH	–5,85	<0,001	спадний	0,0142	статистично значущий, дуже слабкий

Категорія	Z-статистика	p-value	Знак тренду	Cramer's V	Інтерпретація
UA-IDN	-65,60	<0,001	спадний	0,0635	статистично значущий спадний тренд; найсильніший за Z серед категорій, але слабкий за розміром ефекту
UA-WAR	-13,63	<0,001	спадний	0,0329	статистично значущий, слабкий
UA-FRE	-27,33	<0,001	спадний	0,0292	статистично значущий, слабкий

Примітка. p-value подано у скороченому вигляді; Cramer's V розраховано для зв'язку між періодом і бінарною активністю категорії. Усі ефекти залишаються малими або дуже малими за змістовною силою. Джерело 03_trends_cochran_armitage.csv

Результати підтверджують, що в сирих даних усі категорії мають статистично значущий спадний тренд. Найбільша за модулем Z-статистика належить UA-IDN ($Z = -65,60$), далі йдуть CON ($Z = -34,88$) і UA-FRE ($Z = -27,33$). Водночас Cramer's V в усіх випадках залишається малим: навіть для UA-IDN він становить лише 0,0635. Це означає, що статистична значущість насамперед є наслідком великого N, тоді як змістовна сила відмінностей дуже мала.

Спостережувані тренди є відносним зниженням словниково фіксованої активності маркерів і не переносяться напряму на рівень суспільної свідомості (див. застереження в преамбулі 4).

Попарні порівняння періодів і розміри ефекту.

Щоб уточнити, між якими саме періодами відбувалися найпомітніші зміни, виконано попарні порівняння часток активних абзаців. У таблиці Р.1 (Додаток Р) подано три сусідні переходи та порівняння першого й четвертого періодів.

Оскільки всі розрахунки виконуються для пропорцій, як стандартизовану міру розміру ефекту використано Cohen's h (Cohen, 1988). За загальноприйнятою інтерпретацією значення h близько 0,20 відповідає малому ефекту; у цьому корпусі більшість значень є нижчою за цей поріг, що підтверджує діагностичний, а не субстанційно-вимірювальний статус показників.

Попарні порівняння (детальні значення в Табл. Р.1, Додаток Р) виявляють три патерни. Перший: найрізкіший спад UA-IDN — між першим і другим періодами (−4,16 п.п. у сирих даних); надалі категорія знижується значно повільніше. Другий: CON має подібну траєкторію — основне падіння у переході 1991–2003 >> 2004–2013, частковий відскок у третьому періоді, спад у четвертому. Третій: UA-WAR демонструє нелінійний патерн (падіння >> підвищення у 2014–2021 >> спад після 2022), який потребує жанрового контролю в 4.2; сама категорія включає не лише сучасну війну, а й історичні катастрофи та пам'ять про насильство.

Парадокс сирих показників: системне зниження активності категорій.

Головний результат сирого аналізу можна сформулювати як парадокс системного зниження: у прямому агрегуванні знижується активність усіх семи категорій. Змістовно це не виглядає як проста ціннісна трансформація, адже категорії представляють різні, а частково й протилежні мотиваційні полюси. Якщо одночасно падають OPN, CON, TRA, ENH, UA-IDN, UA-WAR і UA-FRE, то принаймні частина ефекту, ймовірно, пов'язана не з односпрямованою зміною цінностей, а з організацією корпусу або способом фіксації маркерів.

Саме цей парадокс емпірично активує методологічне очікування Н4. Якщо жанрова структура корпусу істотно змінюється між періодами, а жанрові блоки відрізняються за щільністю ціннісно маркованої лексики, то сирі тренди можуть

бути зміщені. У такому разі зниження категорії може означати не зниження її присутності всередині однорідного типу дискурсу, а зміну частки жанрів, у яких ця категорія традиційно активується частіше або рідше. Перевірці цієї можливості присвячено підрозділ 4.2.

4.2. Жанрова стандартизація як корекція корпусних зміщень

Жанрова стандартизація застосовується як методична відповідь на проблему неоднорідності корпусу. Її завдання — не “виправити” корпус, а зробити інтерпретацію трендів стійкішою: порівнювати періоди так, ніби жанрова композиція в усіх них була однаковою. Це дозволяє відокремити ту частину зміни, яка зумовлена зміною питомої ваги жанрових блоків, від тієї частини, яка зберігається після контролю жанру і може бути обережно інтерпретована як власне динаміка словникових ціннісних репрезентацій.

Жанрова композиція корпусу за хронологічними зрізами.

Першою підставою для стандартизації є різка зміна жанрової композиції корпусу між першим і четвертим періодами, попередньо описана в Таблиці 2.2 (підрозділ 2.1). У контексті аналізу ціннісних трендів цей зсув набуває іншого змістового значення: якщо в підрозділі 2.1 він описувався як характеристика самого корпусу, то тут він розглядається як потенційне джерело статистичного зміщення сирих трендів активності ціннісних категорій. Тому подальша інтерпретація потребує методу, який дозволяє відокремити жанровий компонент змін від власне ціннісного.

Найсильніший зсув стосується блоку Е: його частка зменшується з 53,35% у 1991–2003 рр. до 15,69% у 2022–2025 рр., тобто на –37,66 п.п (див. Рис. 4.2). Натомість блок С зростає з 4,04% до 35,19% (+31,15 п.п.). Також збільшується частка жанрової белетристики В: з 12,18% до 23,67%. Отже, корпус переходить від домінування історично орієнтованих гібридних і пограничних форм у першому періоді до суттєво більшої присутності фантастики, фентезі та жанрової белетристики у четвертому.

Цей зсув не є технічною помилкою корпусу. Навпаки, він сам по собі може бути соціологічно значущим симптомом трансформації літературного поля (Bourdieu, 1993) й цифрової доступності текстів. Однак для аналізу ціннісних категорій він створює ризик змішування двох різних процесів: зміни власне ціннісної лексики в межах жанрів і зміни питомої ваги жанрів, які мають різні лексико-семантичні профілі. Тому подальший аналіз спирається на стандартизовані показники.



Рис. 4.2. Жанрова композиція україномовного прозового корпусу за чотирма хронологічними зрізами

Метод прямої стандартизації за фіксованою жанровою структурою.

Для контролю жанрового складу застосовано метод прямої стандартизації (Das Gupta, 1993; Kitagawa, 1964). У цьому дослідженні як фіксований стандарт використано середню жанрову структуру всього корпусу: А — 25,03%, В — 18,60%, С — 15,15%, D — 7,72%, Е — 33,49%. Інакше кажучи, для кожного періоду активність категорії в окремих жанрових блоках зважувалася не фактичними частками жанрів у цьому періоді, а спільною для всіх періодів корпусною структурою.

Формально стандартизована активність категорії в періоді розраховується як сума добутків жанрових ваг стандарту на фактичну активність категорії всередині кожного жанрового блоку цього періоду: $pct_std(period, category) = \sum w_genre \times pct(period, genre, category)$. Така процедура не створює нових даних і не змінює результатів кодування. Вона лише змінює спосіб агрегування, щоб різниця між періодами не була механічним наслідком того, що в одному періоді більше історичної прози, а в іншому — більше фантастики чи жанрової белетристики.

Стандартизовані тренди ціннісних категорій.

Після застосування прямої жанрової стандартизації усі сім категорій зберігають нижчі значення у четвертому періоді порівняно з першим, однак масштаб зниження для більшості категорій помітно послаблюється. Таблиця 4.3 подає стандартизовану активність категорій у чотирьох зрізах.

Таблиця 4.3.

Стандартизована активність семи ціннісних категорій за фіксованою жанровою структурою корпусу, % абзаців

Категорія	1991–2003	2004–2013	2014–2021	2022–2025	Δ (4 vs 1), п.п.
OPN	22,17	21,74	22,22	21,37	–0,79
CON	19,07	17,17	17,74	16,62	–2,45
TRA	14,70	13,98	13,96	13,60	–1,10
ENH	16,50	16,15	16,68	15,51	–0,98
UA-IDN	11,11	9,17	8,65	7,86	–3,25
UA-WAR	12,81	11,41	12,59	11,42	–1,39
UA-FRE	7,71	6,70	6,57	6,26	–1,45

Примітка. Стандартизацію виконано за фіксованою середньою жанровою структурою корпусу: $A = 25,03\%$, $B = 18,60\%$, $C = 15,15\%$, $D = 7,72\%$, $E = 33,49\%$.

У стандартизованих даних спадний напрям зберігається для всіх категорій, але масштаб помітно скорочується (Таблиця 4.4, Рис. 4.3 та Рис. 4.4). Найбільше — для UA-IDN (з –6,00 до –3,25 п.п.) і CON (з –4,85 до –2,45 п.п.); істотно слабшає тренд OPN (з –1,72 до –0,79 п.п.) і UA-WAR (з –2,48 до –1,39 п.п.).

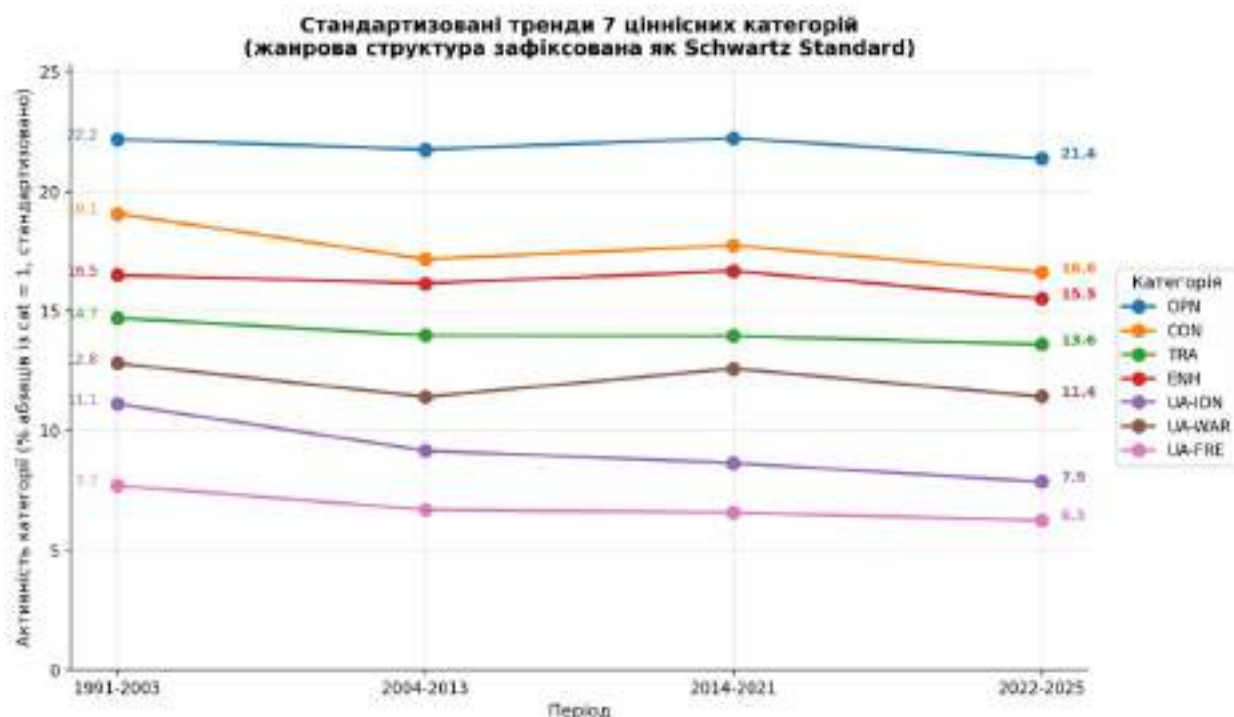


Рис. 4.3. Стандартизовані тренди активності семи ціннісних категорій у чотирьох хронологічних зрізах

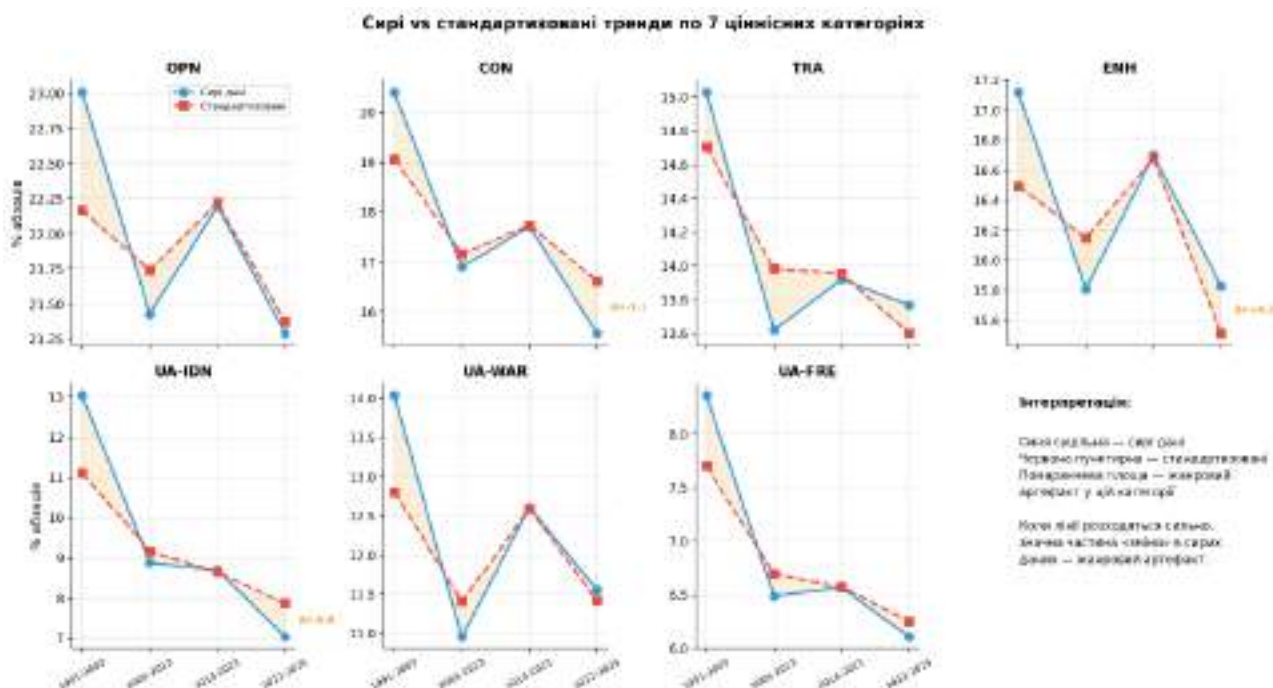


Рис. 4.4. Сирі та стандартизовані тренди ціннісних категорій у порівнянні

Найменший вплив стандартизації спостерігається для TRA та ENH. Для TRA сире зниження становить близько $-1,25$ п.п., стандартизоване — $-1,10$ п.п.; для ENH відповідно $-1,29$ і $-0,98$ п.п. Це означає, що для цих категорій сирий тренд меншою мірою пояснюється жанровим складом корпусу й більшою мірою зберігається після контролю жанру.

Декомпозиція змін: жанровий артефакт і власне динаміка маркерів ціннісних категорій.

Для інтерпретації впливу стандартизації застосовано просту декомпозицію зміни між першим і четвертим періодами: загальна зміна в сирих даних розкладається на жанрову компоненту та власне ціннісну компоненту. У цій логіці Δ_{total} — це різниця між першим і четвертим періодами у сирих даних; $\Delta_{ціннісна}$ — різниця після жанрової стандартизації; $\Delta_{жанрова}$ — залишкова частина, яка зникає або змінюється внаслідок фіксації жанрової структури.

Таблиця 4.4.

Декомпозиція зміни активності категорій між 1991–2003 і 2022–2025 рр.

Категорія	Δ загальна, п.п.	Δ жанрова, п.п.	Δ ціннісна, п.п.	Частка ціннісної компоненти, %	Класифікація
OPN	–1,72	–0,93	–0,79	46,1	тренд частково жанровий артефакт
CON	–4,85	–2,40	–2,45	50,5	тренд частково зумовлений жанровою композицією
TRA	–1,25	–0,15	–1,10	88,0	тренд переважно зберігається після жанрової стандартизації
ENH	–1,29	–0,31	–0,98	76,2	тренд переважно зберігається після жанрової стандартизації
UA-IDN	–6,00	–2,75	–3,25	54,1	тренд частково жанровий артефакт
UA-WAR	–2,48	–1,10	–1,39	55,9	тренд частково зумовлений жанровою композицією
UA-FRE	–2,25	–0,80	–1,45	64,4	тренд частково жанровий артефакт

Примітка. Δ ціннісна відповідає зміні після жанрової стандартизації; Δ жанрова — різниця між сирою і стандартизованою зміною. Частка ціннісної компоненти розрахована за модулем як $|\Delta \text{ ціннісна}| / |\Delta \text{ загальна}| \times 100$. Декомпозицію розраховано за нерозокругленими значеннями часток; тому окремі Δ можуть відрізнятися від різниць, отриманих прямим відніманням округлених відсотків у Таблиці 4.1, на 0,01 п.п. Джерело 04_standardization_decomposition.csv

Декомпозиція показує, що для більшості категорій сирий спад є частково жанровим артефактом. Для OPN жанрова компонента становить $-0,93$ п.п. із загального спаду $-1,72$ п.п.; отже, майже половина сирої зміни зникає після контролю жанру. Подібна ситуація спостерігається для CON, UA-IDN, UA-WAR і UA-FRE. Для UA-IDN жанрова компонента становить $-2,75$ п.п., а стандартизована, тобто власне ціннісна, — $-3,25$ п.п. Це означає, що зниження UA-IDN не можна повністю пояснити зменшенням частки історичної прози, але жанровий зсув суттєво посилював сирий спад.

Графічне представлення розкладу на жанрову і власне ціннісну компоненти для всіх семи категорій наведено на Рисунку С.1 (Додаток С); візуалізація унаочнює, що для TRA та ENH стовпчики «жанрової компоненти» залишаються мінімальними, тоді як для CON, UA-IDN, UA-WAR і UA-FRE жанрова компонента становить майже половину сумарного спаду.

Натомість TRA та ENH належать до категорій, для яких основна частина спадної динаміки зберігається після жанрової стандартизації. Це не означає, що зафіксований спад є “реальною зміною цінностей” у прямому соціологічному сенсі; коректніше говорити, що він меншою мірою пояснюється жанровою рекомпозицією корпусу й більшою мірою пов’язаний із внутрішньожанровою зміною щільності словникових маркерів.

Порівняння Z-статистик Cochran–Armitage до і після стандартизації підтверджує той самий висновок. OPN послаблюється з $Z = -9,02$ до $Z = -3,27$; CON — з $-34,88$ до $-16,78$; UA-WAR — з $-13,63$ до $-5,59$. Для TRA, навпаки, сила тренду практично не змінюється ($-9,79 \gg -9,99$), що вказує на відносну незалежність цього спадного патерну від жанрової композиції. Зведене порівняння абсолютних Z-статистик усіх семи категорій до і після жанрової стандартизації наведено на Рисунку С.2 (Додаток С); рисунок візуально окреслює дві групи: категорії з сильним зменшенням $|Z|$ після стандартизації (CON, UA-IDN, OPN, UA-WAR) і категорії з практично незмінним $|Z|$ (TRA,

ЕНН). Таким чином, Н4 підтверджується не як твердження про повне спотворення всіх трендів, а як твердження про необхідність жанрового контролю для коректної інтерпретації масштабу і частини напрямів змін.

Розбіжність сирих і стандартизованих трендів: інтерпретація на прикладі OPN.

Категорія OPN є зручною для демонстрації значення жанрової стандартизації. У сирих даних OPN знижується з 23,01% до 21,29%, тобто на $-1,72$ п.п. Після стандартизації падіння зменшується до $-0,79$ п.п. При цьому напрям тренду за обраним корпусним стандартом не змінюється: OPN усе ще має спадний знак, але його сила значно слабша. Z-статистика зменшується з $-9,02$ до $-3,27$, тобто сирий тренд був суттєво підсилений жанровим складом корпусу.

Цей приклад принциповий методологічно. Якби аналіз обмежувався сирими даними, дослідник міг би зробити висновок про помітне зниження відкритості до змін у художньому дискурсі. Після стандартизації такий висновок має бути значно обережнішим: частина зниження пов'язана зі зміною жанрової композиції, насамперед із різким скороченням блоку Е і зростанням блоку С. Отже, коректне формулювання полягає не в тому, що OPN “різко падає”, а в тому, що за корпусним стандартом зберігається слабкий спадний тренд OPN, істотно послаблений після контролю жанру.

У підсумку приклад OPN показує, чому результати 4.1 не можуть бути фінальними: сирі тренди мають діагностичну функцію, але змістовна інтерпретація має спиратися на стандартизовані показники та декомпозицію (Рис. 4.5). У підрозділі 4.3 саме стандартизовані результати використовуються як основа для перевірки теоретично мотивованих очікувань.

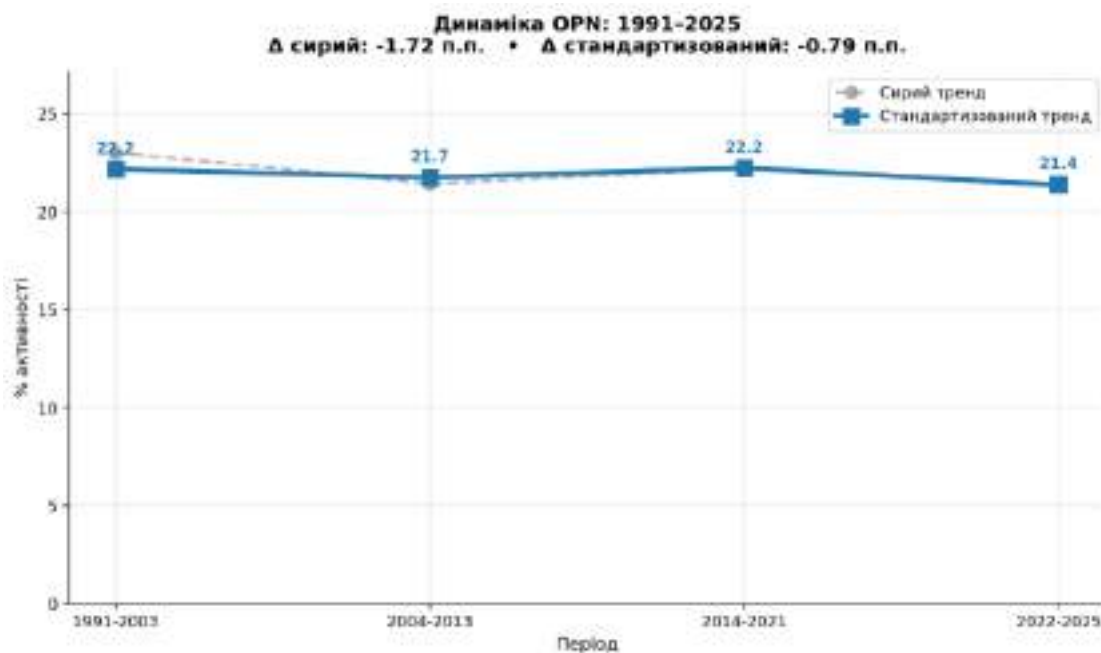


Рис. 4.5. Сирий і стандартизований тренд OPN: вплив жанрової корекції

4.3. Перевірка теоретичних гіпотез у світлі емпіричних результатів

У цьому підрозділі стандартизовані результати зіставляються з трьома очікуваннями, сформульованими в підрозділі 1.5. Інтерпретація здійснюється у двох площинах: по-перше, на агрегатному рівні всього корпусу; по-друге, з урахуванням жанрово-стратифікованих результатів, отриманих з тієї самої таблиці активності за комбінаціями (період $\times$ жанровий блок $\times$ категорія), що використовувалась для жанрової стандартизації в підрозділі 4.2 ($n = 140$ рядків).

H1: динаміка UA-IDN у хронологічних зрізах

H1 передбачала, що активність категорії UA-IDN має демонструвати локальні підвищення після ключових громадянсько-мобілізаційних подій: у

періоді 2004–2013 рр. після Помаранчевої революції та в періоді 2014–2021 рр. після Революції Гідності й початку війни на Донбасі. У сирих даних цього патерну не зафіксовано: UA-IDN послідовно знижується від 13,04% до 8,88%, далі до 8,70% і 7,04%.

Після жанрової стандартизації картина змінюється за масштабом, але не за напрямом: 11,11% у 1991–2003 рр., 9,17% у 2004–2013 рр., 8,65% у 2014–2021 рр. і 7,86% у 2022–2025 рр. Отже, очікуваних локальних піків не виявлено. Н1 на агрегатному рівні не підтверджується. Водночас наявні стратифіковані результати уточнюють агрегатну картину. У блоці А (реалістична та сучасна проза) активність UA-IDN не падає монотонно, а демонструє локальний пік саме в зрізі 2004–2013 рр.: 8,02% >> **9,57%** >> 8,98% >> 6,32%. Це частково підтверджує Н1s у тій частині, що стосується пост-помаранчевого піку в сучасних жанрах: реалістична проза 2004–2013 рр. дійсно активніше експлуатує сучасний шар української ідентифікаційної лексики, ніж у попередньому або наступному зрізі. На агрегатному рівні цей пік розчиняється через одночасне різке падіння UA-IDN у блоці Е (з 17,56% у 1991–2003 рр. до 10,51% у 2004–2013 рр.) і скорочення частки самого блоку Е з 53,35% до 28,59%. Отже, агрегатне спростування Н1 не можна пояснити лише жанровою рекомпозицією корпусу. Жанровий зсув суттєво посилює спад UA-IDN, але після стандартизації спадна динаміка зберігається. Тому коректний висновок полягає в тому, що Н1 не підтверджується на рівні всього корпусу, проте Н1s частково підтримується для блоку А, де справді зафіксовано локальне підвищення UA-IDN у періоді 2004–2013 рр.

Цей результат потребує обережної інтерпретації. Він не означає зниження значущості української ідентичності в суспільстві після 2004 чи 2014 рр. Скоріше він показує, що на рівні використаного словника й агрегованого корпусу маркери UA-IDN мають вищу активність у першому зрізі, де жанрово й тематично сильніше представлена історична та ретроспективна проза. Після

стандартизації спад UA-IDN зменшується з $-6,00$ до $-3,25$ п.п., але не зникає. Це дає підстави сформулювати окремий методичний висновок: словник UA-IDN, імовірно, краще фіксує історико-культурну й ретроспективну ідентичнісну лексику, ніж сучасніші форми ідентифікації 2014–2025 рр., які можуть виражатися через непрямі, контекстні або сюжетно-наративні маркери.

H2: зміни ENH і проблема очікуваного монотонного зростання

H2 передбачала, що категорія ENH / Self-Enhancement, яка агрегує досягнення і владу, має демонструвати тенденцію до зростання впродовж 1991–2025 рр. як індикатор постсоціалістичного модернізаційного зсуву до індивідуалістичних орієнтацій. На агрегатному рівні це очікування не підтверджується. У сирих даних ENH змінюється нелінійно: $17,12\% \gg 15,81\% \gg 16,70\% \gg 15,83\%$. У стандартизованих даних патерн подібний: $16,50\% \gg 16,15\% \gg 16,68\% \gg 15,51\%$. Загальна зміна між першим і четвертим періодами становить $-0,98$ п.п.

Отже, H2 як очікування монотонного зростання ENH на рівні всього корпусу спростовується. Водночас цей результат не обов'язково суперечить ширшій логіці модернізаційних теорій (Inglehart & Welzel, 2005), якщо врахувати кризовий характер української траєкторії. Повномасштабна війна, екзистенційні загрози, переорієнтація на безпеку та колективну стійкість можуть блокувати або послаблювати лінійну динаміку самоствердження в агрегованому літературному дискурсі (Welzel, 2013).

Водночас наявні стратифіковані результати уточнюють картину. Для блоку В, тобто жанрової белетристики, зафіксовано зростання ENH з $15,64\%$ до $19,31\%$, тобто на $+3,67$ п.п. Це означає, що H2s частково підтримується саме в тому сегменті корпусу, де очікувано сильніше представлені моделі особистого успіху, конкуренції, статусу й життєвої самореалізації. Блок D не демонструє

аналогічного підтвердження H2s: активність ENH у дитячо-підлітковій прозі змінюється нелінійно і не завершується зростанням у четвертому зрізі. Тому H2s підтримується не для всієї передбаченої групи жанрів, а насамперед для блоку В. Отже, коректне формулювання висновку таке: H2 не підтверджується агрегатно, але має часткову жанрово-стратифіковану підтримку в блоці В.

H3: TRA після 2022 р. і проблема очікуваного реактивного зростання.

H3 передбачала, що активність TRA / Self-Transcendence має зрости у періоді 2022–2025 рр. як реакція на воєнний контекст: через посилення мотивів солідарності, турботи, взаємодопомоги, жертвовності та універсалістських орієнтацій. На агрегатному рівні це очікування також не підтверджується. У сирих даних TRA у 2014–2021 рр. становить 13,92%, а у 2022–2025 рр. — 13,77%, тобто змінюється на –0,15 п.п. У стандартизованих даних спад між третім і четвертим періодами більший: 13,96% >> 13,60%, або –0,36 п.п.

Це означає, що H3 у формі загального агрегатного очікування про післявоєнний сплеск самотрансцендентності не підтверджується. У блоці А (де теоретично очікувалося найбільш виразне зростання TRA згідно з H3s) динаміка становить 13,93% >> 13,57% >> 15,17% >> 14,03%. Це означає, що максимум TRA в реалістичній і сучасній прозі припадає не на четвертий зріз, як прогнозувала H3s, а на третій — період між Революцією Гідності та повномасштабною війною (15,17% у 2014–2021 рр.). У зрізі 2022–2025 рр. TRA в блоці А знижується до 14,03% (–1,14 п.п. порівняно з третім зрізом), хоча залишається вищою за рівні першого і другого зрізів. Водночас у блоках В і D спостерігається інший патерн: у жанровій белетристиці TRA зростає з 11,65% у 1991–2003 рр. до 15,14% у 2022–2025 рр., а в дитячо-підлітковій прозі — з 10,08% до 12,89%. Отже, H3s не підтверджується в первісно очікуваному вигляді для блоку А, але частково підтримується для блоків В і D. Це свідчить, що

самотрансцендентні маркери після 2022 р. активізуються не в усьому корпусі й не насамперед у реалістичній прозі, а в окремих жанрових сегментах. Це **істотно ускладнює інтерпретацію НЗs**: реактивне посилення самотрансцендентних маркерів у воєнно-орієнтованій прозі дійсно зафіксовано, але воно припадає на період Донбаської війни 2014–2021 рр., а не на постфевральський період. Можливе пояснення — **аксіологічний лаг** (підрозділ 1.2) у його літературному варіанті: художня рефлексія повномасштабної війни 2022 р. на момент формування корпусу (станом на 2025 р.) перебуває ще в ранній фазі і не встигла перейти у щільно лексикалізовану форму, а вже відрефлексований досвід донбаського періоду 2014–2021 рр. дав достатній часовий запас для нарощування лексично-маркованих солідарнісно-універсалістських мотивів. Тобто пік TRA в блоці А фіксується не на тому етапі суспільного досвіду, який його мав би теоретично породити, а на наступному — коли суспільство вже встигло цей досвід художньо опрацювати. Окремо, як зазначено в 3.1, TRA серед усіх семи категорій має найвищий α Криппендорфа в LLM-кодуванні (0,708), що дає цьому результату додаткову формальну надійність на лексичному рівні.

Такий результат є важливим змістовно. Він показує, що агреговані корпусні показники можуть приховувати локальні жанрові зміни. Самотрансцендентність після 2022 р. не стає домінантним макротрендом у всьому корпусі, але її маркери активізуються в тих сегментах, які можуть бути більш чутливими до масового читання, соціалізації та емоційної переробки воєнного досвіду. Додатково зафіксоване зростання UA-WAR у блоці В між третім і четвертим зрізами (2014–2021 >> 2022–2025) — з 9,83% до 12,22% (+2,39 п.п.) — підтримує інтерпретацію про реакцію жанрової літератури на повномасштабну війну. При порівнянні першого і четвертого зрізів зростання UA-WAR у блоці В становить +2,09 п.п. (10,13% >> 12,22%), що менш виразне, але збігається за напрямом.

Підсумково H1, H2 і H3 не підтверджуються на агрегатному рівні всього корпусу. Однак H1s, H2s і H3s мають часткову жанрово-стратифіковану підтримку. Це виводить дослідження від простої логіки «гіпотеза підтверджена / спростована» до більш соціологічно продуктивного висновку: динаміка маркерів ціннісних категорій художнього дискурсу є нерівномірною, жанрово опосередкованою і залежною від того, який сегмент літературного поля аналізується.

4.4. Якісна інтерпретація показових і контрастних кейсів

Кількісні результати 4.1–4.3, навіть після жанрової стандартизації, не реконструюють самостійно ціннісну семантику художнього дискурсу. Як зафіксовано в підрозділі 2.2 (Межі 2, 4), автоматизоване кодування реєструє наявність лексичних або n-грамних маркерів, але не завжди здатне відрізнити пряму авторську позицію від мовлення персонажа, іронічної дистанції, історичної цитати чи наративно вбудованої ціннісної опозиції. У цьому підрозділі стандартизовані макротренди доповнюються цільовим якісним аналізом семи фрагментів, відібраних з трьох V1-метаданих файлів за принципом стратифікованої максимальної інформативності (Krippendorff, 2019; Grimmer et al., 2022). Три кейси подано розгорнуто в основному тексті — два показові (постпомаранчевий пік UA-IDN у блоці А та пік TRA у блоці А в добу Донбаської війни) і один контрастний (систематична недооцінка сучасного ідентичнісного дискурсу). Чотири додаткові кейси — два показові (пік UA-IDN у блоці Е у 1991–2003 рр. та зростання ENH у блоці В після 2022 р.) і два контрастні (повторний випадок недооцінки UA-IDN на матеріалі прози В. Рафєєнка і повна нульова автоактивація на матеріалі архаїчно-стилізованої прози В. Лиса) — розгорнуто в Додатку П.

Кожен фрагмент аналізується на трьох рівнях: **лексико-семантичному** (які слова або словосполучення активують категорію), **наративному** (хто є носієм ціннісної позиції — наратор, персонаж, рефлексуючий «я», іронічний голос) і **дискурсивному** (як фрагмент функціонує у ширшому соціокультурному контексті свого періоду й жанру).

Кейси, що ілюструють виявлені тренди.

Кейс 1. Костенко Ліна. «Записки українського самашедшого», (2010) (блок А, 2004–2013). Наративний абзац із 93 слів (автокод: OPN + UA-IDN):

«Тим часом новий скандал. Без скандалу ми не можемо. То касетний, то кольчужний, а це уже зерновий... Так що всупереч нашій національній символіці — синє небо над золотим полем — імпортуємо пшеницю з країни, символ якої — кленовий лист...».

На лексико-семантичному рівні UA-IDN активується через пряме посилення на національну символіку («синє небо над золотим полем»), OPN — через лексику новизни і відкритості до зовнішнього («новий скандал», «імпортуємо»). На наративному рівні фрагмент організовано як іронічну каталогізацію (касетний >> кольчужний >> зерновий), де голос першоособового наратора створює колективне «ми» через спільне розчарування; національний символ розгортається не як декларація, а як риторично-критичний фон. На дискурсивному рівні фрагмент є **прямою емпіричною ілюстрацією часткового підтвердження H1s** у підрозділі 4.3: пост-помаранчевий пік активності UA-IDN у блоці А (9,57% у 2004–2013 рр. проти 8,02% у попередньому зрізі) формується саме такими саморефлексивно-критичними апеляціями до національних символів, а не їхнім декларативним утвердженням. Інтертекстуальна апеляція в самій назві твору до гоголівських «Записок сумасшедшего» (з доданим *українського*) є додатковим жестом ідентифікаційного позиціонування у

відмінному від російської літературної традиції просторі. Водночас цей приклад демонструє не лише можливість виявлення OPN-маркерів, а й обмеження словникової процедури: лексеми на кшталт «новий» або «імпортуємо» можуть активувати категорію відкритості до змін формально, хоча в конкретному контексті вони не завжди позначають автономію, самостійність чи готовність до інновації у шварцівському сенсі. Тому такі випадки потребують якісної перевірки.

Кейс 2. Лис Володимир. «Обітниця», (2019) (блок А, 2014–2021). Діалоговий фрагмент із 32 слів (автокод: TRA з найвищою серед усіх кандидатів інтенсивністю 6,0):

«— Колись би сказали — кінік, — Андрій цокнувся склянкою з Богдановою. — Хоч ті філософи були ліпші, ніж ми з тобою. І ти ліпший за мене. А мо', й за себе.»

На лексико-семантичному рівні TRA активується через триразове повторення компаратива *ліпший* — маркера моральної оцінки інших і самооцінки через інших, що належить до ядра шварцівської доброзичливості й універсалізму. На наративному рівні фрагмент організовано як інтимне визнання моральної переваги співрозмовника («ти ліпший за мене»), яке через подвоєння («а мо', й за себе») зміщується у бік моральної саморефлексії. На дискурсивному рівні фрагмент **підтверджує жанрово-внутрішню підтримку прогнозів 4.3** у блоці А в зрізі 2014–2021 рр., де TRA сягає піка 15,17%: пост-Майданна реалістична проза в період Донбаської війни актуалізує лексично-марковану самотрансцендентну рамку через інтимні діалоги моральної оцінки, а не через прямі солідарнісні декларації. Принципово, що цей пік припадає не на четвертий зріз, як прогнозувала НЗ, а саме на 2014–2021 рр., — що узгоджується з висновком 4.3 про **вибірковий, а не реактивний** характер посилення TRA в блоці А. Повторення компаратива «ліпший» саме по собі не є достатнім маркером самотрансцендентності. Його можна інтерпретувати як сигнал TRA

лише у зв'язку з контекстом моральної оцінки іншого, орієнтацією на турботу, підтримку або визнання людської гідності. Тому цей кейс демонструє необхідність контекстної перевірки автоматично активованих категорій.

Кейс, що ускладнює або суперечить кількісним трендам.

Кейс 3. Криштопа О., Каліновська О. «Про мертвих, живих і ненароджених» (2016) (блок А, 2014–2021). Наративний фрагмент із 83 слів (автокод: OPN + ENH + UA-WAR, але не UA-IDN):

«Уперше я приїхав на Донбас 2004-го. Мене питали, чи я із Західної України — через українську мову, а я відповідав, що ні, місцевий. І частина правди в цьому була. Мій дід, Віктор Криштопа, народився на Донбасі, неподалік Волновахи, і з дитинства розмовляв українською. Після закінчення Київського університету радянська влада відправила його на крайній “дикий” захід — до Івано-Франківська, перевиховувати “бандерівців”... Постійні журналістські відрядження, які руйнували журналістські міфи...»

Цей кейс — **пряма емпірична ілюстрація проблеми, виявленої у підрозділі 3.2** ($F1 = 0$ для UA-IDN проти всіх трьох gold standards). На лексико-семантичному рівні фрагмент містить надзвичайно щільне ідентичнісне лексичне поле (українська мова, Західна Україна, Київський університет, бандерівці, Донбас, Івано-Франківськ), однак UA-IDN автоматично не активується; натомість UA-WAR активується через географічну прив'язку «Донбас». На наративному рівні першоособова мемуарно-журналістська рефлексія структурує родинну історію як зворотний рух між Заходом і Сходом, делегітимізуючи спрощену регіональну стереотипізацію. На дискурсивному рівні інтертекстуальна назва книги Криштопи — пряма апеляція до Шевченка («І мертвим, і живим, і ненародженим землякам моїм...») — підкреслює саме той сучасний шар ідентичнісної лексики, до якого поточний словник UA-IDN, налаштований переважно на ранньонезалежницький постколоніальний шар,

виявляється нечутливим. Цей кейс емпірично обґрунтовує сформульовану в підрозділі 4.6 потребу першочергової ревізії словника UA-IDN.

Підсумок 4.4 Сім розглянутих кейсів сукупно демонструють обидва типи операційного обмеження автоматизованого підходу: переоцінку маркерної щільності (**Кейси П.1 та П.2** — на матеріалі І. Білика та А. Новіка — словник реєструє ENH і UA-IDN на лексиці, що в контексті жанру або періоду має ослаблене ціннісне навантаження), систематичну недооцінку сучасного ідентичнісного дискурсу (**Кейс 3** в основному тексті на матеріалі О. Криштопи та **Кейс П.3** на матеріалі В. Рафєєнка — два випадки нульової активації UA-IDN при дуже високій змістовій ідентичнісній насиченості) і повну невидимість ціннісного шару при роботі з нетиповою стилістичною модальністю (**Кейс П.4** на матеріалі В. Лиса «Вифлеєм», Додаток П). Водночас два показові кейси основного тексту (**Кейс 1** Костенко, **Кейс 2** Лис «Обітниця») разом із Кейсом П.1 (Білик) яскраво підтверджують жанрово-стратифіковані висновки підрозділу 4.3 — пост-помаранчевий пік UA-IDN у реалістичній прозі та доба Донбаської війни як період посилення TRA в блоці А — і одночасно ілюструють, чому жанрова стандартизація є необхідною: ці локальні підтвердження зникають на агрегатному рівні через одночасні зміни жанрової композиції корпусу.

4.5. Зіставлення з даними WVS, EVS і національними опитуваннями: діагностична перевірка

Корпусний підхід і опитувальні дослідження ціннісних орієнтацій вимірюють принципово різні рівні ціннісної системи. Корпус фіксує режим лексичної репрезентації цінностей у художньому дискурсі — частоту й виразність їхнього лексичного маркування у текстах. World Values Survey (Haerpfer et al., 2022) і European Values Study (2022) вимірюють декларативно-

самозвітний рівень — те, що респонденти готові формулювати в інтерв'ю на момент проведення хвили. Розходження між цими рівнями не свідчить про помилковість одного з них; воно є аналітично значущим індикатором того, що одна й та сама ціннісна сфера може по-різному проявлятися в декларативно-опитувальному та художньо-дискурсивному режимах. Тому зіставлення корпусу з WVS і EVS у цьому підрозділі виконує діагностичну, а не верифікаційну функцію, і його результати слугують одним із вхідних даних для узагальненої інтерпретації у підрозділі 4.6.

Для зіставлення кожна із семи категорій кодбука співвіднесена з основною WVS-змінною та її EVS-аналогом за принципом концептуальної близькості; повна карта операціональних відповідностей між сімома категоріями кодбука й шкалами WVS/EVS (з main- і secondary-змінними, EVS-відповідниками і напрямком нормалізації) подана у Таблиці Т.1 Додатку Т. Часова відповідність хвиль, методологічні розвилки розрахунку та опорні точки національних опитувань 2022–2025 рр. також винесені до Додатку Т. У основному тексті подано підсумкове зіставлення напрямків трендів між першою і останньою доступною точкою для України в кожній з оптик. Дані подані в таблиці 4.5.

Таблиця 4.5.

Зіставлення напрямків трендів: корпус (стандартизована активність), WVS, EVS

Категорія	Корпус, Δ п.п.	WVS, Δ норм. 0–100	EVS, Δ норм. 0–100	Тип збігу
OPN	–0,8	Y003_simple +2,4	Y003_simple +7,5	DIVERGENT
CON	–2,4	Y010 +1,6	F063 +7,3	DIVERGENT
TRA	–1,1	A165 –0,1	A165 +1,0	FLAT
ENH	–1,0	Y001 –5,0	E035 –9,2	CONVERGENT
UA-IDN	–3,2	G006 +12,7	G006 +15,7	DIVERGENT
UA-WAR	–1,4	E069_02 +3,6	E069_02 +4,6	DIVERGENT

UA-FRE	-1,5	Y024 +3,2	E117 +6,2	DIVERGENT
--------	------	-----------	-----------	-----------

Примітка. Корпусні значення — зміна стандартизованої частки активних абзаців між зрізом 1991–2003 і зрізом 2022–2025 рр. (за результатами підрозділу 4.2). WVS — зміна нормалізованого до шкали 0–100 значення основної змінної між Wave 3 (1996) і Wave 7 (2020) для України. EVS — зміна нормалізованого значення між Wave 3 (1999) і Wave 5 (2017) для України. Технічні характеристики залучених хвиль обох архівів (рік, обсяг вибірки, метод збору даних) подано у Таблиці Т.3 (характеристики хвиль) Додатку Т.

Зіставлення виявляє чотири змістовні патерни. Перший — ENH як контрольна точка тріангуляції: усі три незалежні оптики синхронно фіксують зниження маркерів та індикаторів самоствердження. Там, де концептуальна близькість між корпусним маркером і шкалою опитування максимальна (досягнення, статус, економічна мотивація), оптики сходяться. Це свідчить, що процедура зіставлення здатна фіксувати конвергенцію там, де концептуальна близькість між корпусними маркерами й опитувальними індикаторами є відносно високою. Відповідно, розходження в інших категоріях не можна автоматично пояснити лише технічною несумірністю шкал. Другий — TRA як синхронна стабільність: ні корпус, ні два опитування не фіксують зростання межособистісної довіри упродовж 1996–2020 рр. Двома незалежними консорціумами шкала A165 («Most people can be trusted») дає в Україні практично однакові значення в усіх доступних точках (близько 26–31% згоди), що методично підсилює корпусну знахідку про відсутність зростання маркерів TRA на агрегатному рівні всупереч поширеному у публіцистиці тезису про «сплеск солідарності» після 2014 і 2022 рр.

Третій патерн — систематична дивергенція для п'яти категорій (OPN, CON, UA-IDN, UA-WAR, UA-FRE). WVS і EVS згідно фіксують зростання декларативних показників, тоді як корпус — зниження активності лексичних маркерів. Особливо виразним є випадок UA-IDN: у 2020 р. WVS G006 дає для України 72,7 пункту, EVS G006 — 72,8 пункту, тобто два незалежні консорціуми з різними методиками вибірки зафіксували практично ідентичне значення з різницею 0,1 пункту, тоді як корпусна стандартизована активність UA-IDN падає

з 11,1% у першому зрізі до 7,9% у четвертому. Графічне представлення розширеного часового ряду (корпус, WVS- і EVS-точки для України за 1996–2023 рр.) та зіставлення динаміки сімох ціннісних категорій корпусу з даними WVS подано на Рисунках Т.1 і Т.2 Додатку Т; внутрішню узгодженість main- і secondary-показників WVS усередині кожної категорії візуалізовано на Рисунку Т.3. Збіг WVS і EVS у всіх п'яти DIVERGENT-випадках є важливим методологічним аргументом: жодного разу два архіви не дали взаємно протилежних напрямків, тому дивергенція не пояснюється однією помилковою опитувальною оптикою.

Четвертий патерн — структурне розходження між корпусним і опитувальним режимами для більшості категорій, що підтримує тезис про зсув режиму репрезентації цінності. Цей патерн можна інтерпретувати як зсув режиму репрезентації: частина цінностей, що в ранніх жанрових сегментах корпусу прямо лексично маркувалася, у пізніших текстах частіше переходить у фонові, сюжетні або наративно опосередковані форми. Саме тому зниження активності маркерів не слід автоматично тлумачити як зниження суспільної значущості відповідних цінностей. Контекст 2022–2025 рр., не покритий хвилями WVS і EVS, реконструюється за національними опитуваннями (Razumkov Centre, 2024; Соціологічна група «Рейтинг» & Gallup International, 2023; Фонд «Демократичні ініціативи» імені Ілька Кучеріва & KMIC, 2025) і узгоджується з WVS/EVS-трендами до 2020 р. для трьох з чотирьох доступних категорій: UA-IDN, UA-WAR і UA-FRE. Опорні точки 2022–2025 рр. подано у Додатку Т, п. Т.5.

Зведений емпіричний результат тріангуляції такий: одна категорія дає синхронний рух (ENH — CONVERGENT), одна — синхронну стабільність (TRA — FLAT), п'ять категорій демонструють стійку дивергенцію корпусу і опитувальних оптик (OPN, CON, UA-IDN, UA-WAR, UA-FRE — DIVERGENT). Стійкість висновку перевірено через альтернативний композитний WVS-

показник, що поєднує main- і secondary-змінні всередині кожної категорії (Таблиця Т.2 Додатку Т): композитна спрямованість збігається з main-only-результатом для шести категорій із семи; розбіжність зафіксовано лише для CON, де main-only дає DIVERGENT, а композит — CONVERGENT, що відображає реальну внутрішню неузгодженість концепту збереження. Цей патерн не підтверджує і не спростовує корпусних результатів як таких, а вказує на дисоціацію режимів репрезентації цінностей у художньому дискурсі і у декларативно-самозвітному режимі — феномен, попередньо концептуалізований у підрозділі 1.3. Подальша узагальнена інтерпретація трансформації ціннісних репрезентацій (підрозділ 4.6) будується з урахуванням як корпусних трендів (підрозділи 4.1–4.4), так і виявленої тут дисоціації.

4.6. Висновки до розділу 4

Узагальнення результатів аналізу — корпусних трендів (4.1), жанрової стандартизації (4.2), перевірки теоретичних гіпотез (4.3), якісної інтерпретації показових і контрастних кейсів (4.4) і діагностичного зіставлення з даними WVS, EVS та національних опитувань (4.5) — дозволяє сформулювати узагальнену картину динаміки ціннісних репрезентацій в україномовному художньому дискурсі 1991–2025 рр., зафіксувати методологічні обмеження її інтерпретації та окреслити перспективи подальшої роботи.

На рівні сирих агрегатів корпус демонструє статистично значущий спадний тренд активності всіх семи категорій між першим і четвертим періодами; водночас Cramer's $V \leq 0,07$ і Cohen's h , нижчий за поріг малого ефекту, свідчать про надстатистичну значущість за малою змістовною силою, типову для корпусу обсягом понад 1,3 млн абзаців. Жанрова стандартизація істотно послаблює цей тренд для п'яти категорій (OPN, CON, UA-FRE, UA-WAR, UA-

IDN). Приблизно від третини до половини сирого спаду пояснюється не зміною ціннісної маркерної щільності, а жанровою рекомпозицією корпусу. Цей результат підтверджує методологічне очікування Н4 і обґрунтовує перехід від сирих показників до стандартизованих як основи змістовної інтерпретації. На відміну від решти категорій, TRA та ENH виявляються особливо стійкими до жанрової корекції: після стандартизації понад 75 % спостереженого спаду залишається власне ціннісною компонентою.

Агрегований рівень водночас приховує суттєві жанрові відмінності. Гіпотези Н1, Н2 і Н3 не підтверджуються для корпусу в цілому, проте знаходять часткове підтвердження в окремих жанрових блоках: зростання ENH локалізується в жанровій белетристиці (блок В), а зростання TRA — у блоках В і D (дитяча та підліткова проза), де моделі моральної соціалізації та взаємодопомоги активізуються переважно в текстах, орієнтованих на формування ціннісних орієнтирів аудиторії. Поширений у публічному дискурсі тезис про посилення солідарності після 2014 і 2022 рр. не знаходить прямого підтвердження на рівні агрегованих маркерів TRA та опитувального індикатора міжособистісної довіри A165, що не заперечує можливого зростання інших форм воєнної солідарності, але показує їх нетотожність загальному показнику довіри.

Найбільш обережна змістовна інтерпретація сукупних результатів полягає в тому, що отримані дані можуть свідчити не про послаблення відповідних орієнтацій, а про зміну способу їхньої текстової репрезентації — від експліцитного лексичного маркування в ранніх жанрових сегментах корпусу до нелексикалізованих форм у пізніших. Діагностичне зіставлення з опитувальними оптиками (підрозділ 4.5) надає цьому тлумаченню зовнішню підтримку: ENH функціонує як CONVERGENT-точка триангуляції (синхронне зниження в корпусі, WVS і EVS), TRA — як FLAT (синхронна стабільність), а п'ять із семи категорій демонструють стійку дивергенцію між декларативним зростанням і зниженням лексичних маркерів. Такий патерн доцільно інтерпретувати не як

пряме підтвердження або спростування реальних ціннісних змін, а як підставу для робочої концептуалізації дисоціації режимів репрезентації цінностей. Йдеться про можливе розходження між декларативно-самозвітним режимом, у якому цінності фіксуються через відповіді респондентів на стандартизовані запитання, і художньо-дискурсивним режимом, у якому вони можуть переходити від прямого лексичного маркування до наративно-фонових, сюжетно-композиційних, символічних або інтертекстуальних форм. У цьому сенсі поняття дисоціації режимів репрезентації виконує не доказову, а інтерпретативно-методологічну функцію.

Сформульована інтерпретація має кілька принципових обмежень. **Перше** пов'язане з природою автоматизованого словникового кодування: метод фіксує наявність маркерів, але не завжди розрізняє контекст, носія висловлювання, іронічний режим або цитатність; тому отримані відсотки є індикаторами активності маркерів, а не прямими вимірами ціннісних орієнтацій. **Друге** стосується валідності категорій: зафіксований у Розділі 3 розрив між автоматизованим кодуванням і людськими gold standards ($F1 = 0,076$ проти Human-gold; систематичне переоцінювання у $3\text{--}10\times$) не скасовує можливості макрорівневого порівняння, але вимагає інтерпретації результатів як патернів, а не точних вимірів абсолютної частоти цінностей. Окремо це стосується категорії UA-IDN, на яку спирається найвиразніший емпіричний приклад дисоціації: її спад може частково відображати не реальне послаблення ідентичнісних репрезентацій, а недостатню чутливість словника до новіших форм української ідентичності, що виражаються через практики спротиву, громадянської суб'єктності або мовчазну нормалізацію українськості, а не лише через прямі мовно-культурно-історичні маркери. Тому приклад UA-IDN слід використовувати як найвиразніший індикатор проблеми, але не як самодостатній доказ дисоціації: його переконливість залежить від подальшого уточнення словника, включення нових форм громадянсько-політичної та повсякденної

українськості, а також від ширшої якісної перевірки нелексикалізованих форм ідентичнісної репрезентації. Це не знімає основного аргументу — його методологічно підсилює функціонування ENH як CONVERGENT-точки і TRA як FLAT-точки тріангуляції, що знижує ймовірність систематичного зміщувального артефакту, — але вимагає обережності у формулюваннях і паралельної подальшої ревізії словника.

Третє обмеження стосується корпусу: попри значний обсяг, його жанрова структура нерівномірна, а четвертий період 2022–2025 рр. коротший і залежний від доступності нових цифрових текстів; жанрова стандартизація зменшує цей ризик, але не усуває повністю ефект доступності, авторського складу й видавничої селекції. Четверте — обсяг якісного шару: аналіз семи фрагментів виконує верифікаційно-інтерпретативну функцію і не замінює систематичної якісної вибірки; для статистичної оцінки узгодженості кількісних та якісних висновків необхідне розширення до 30–50 фрагментів зі стратифікацією за матрицею період × жанр.

Виявлені обмеження окреслюють напрями подальшого розвитку запропонованого підходу. Першочерговою є структурна ревізія словника UA-IDN із залученням сучасного шару української ідентичнісної лексики 2014–2025 рр., а також аналогічна, менш масштабна ревізія словників UA-FRE і UA-WAR. Другий напрям — поглиблення жанрово-стратифікованого аналізу з повною статистичною перевіркою гіпотез H1s–H3s у кожному жанровому блоці. Третій — розширення інструментарію на маркери нелексикалізованих режимів репрезентації: сюжетні структури морального вибору, фокалізацію, інтертекстуальні відсилання та символічні структури. Систематичне розрізнення режимів репрезентації цінностей у художньому дискурсі може бути запропоновано як перспективний напрям соціологічного аналізу культурного поля, однак його подальше обґрунтування потребує окремого дослідження.

ВИСНОВКИ

У магістерській дисертації вирішено наукове завдання соціологічної реконструкції трансформацій репрезентацій ціннісних орієнтацій у дискурсі україномовної художньої прози 1991–2025 рр. Розроблено й апробовано інтегрований текстоцентричний підхід, що поєднує автоматизоване аксіологічне кодування великого корпусу художніх текстів, гібридну процедуру валідації, жанрову стандартизацію, якісну інтерпретацію показових кейсів та діагностичне зіставлення з даними World Values Survey, European Values Study й національних опитувань 2022–2025 рр. Підхід запропоновано як комплементарний до опитувальних методів інструмент ретроспективного аналізу ціннісних репрезентацій у завершених історичних періодах, недоступних для прямого опитування, але не як заміну стандартизованим соціологічним вимірюванням.

Прийнято соціологічний ракурс як теоретичне ядро дослідження, що трактує художній дискурс як простір колективно поділюваних ціннісних репрезентацій. Розмежовано чотири рівні аксіологічного аналізу — цінність, ціннісну орієнтацію, ціннісну систему, ціннісну установку — і чотири рівні їх репрезентації в тексті: лексико-семантичний, наративний, дискурсивний та інтертекстуальний. Соціально-історичну рамку українського суспільства операціоналізовано через чотири хронологічні зрізи (1991–2003, 2004–2013, 2014–2021, 2022–2025), що відповідають фазам пострадянської аномії, ціннісної поляризації і громадянської мобілізації, прискореного перевизначення ідентичностей в умовах війни на Донбасі та екзистенційної ціннісної реконструкції.

Сформовано аналітичний корпус із 664 творів 235 авторів обсягом 1 375 333 абзаців. Розроблено категоріальну систему, яка поєднує чотири агреговані виміри моделі Ш. Шварца (OPN, CON, TRA, ENH) з трьома контекстно-

ціннісними категоріями українського художнього дискурсу (UA-IDN, UA-WAR, UA-FRE), що функціонують як аналітичні рамки ідентичності, травми пам'яті та свободи й колективної суб'єктності. Валідацію автоматизованого інструментарію проведено за двома лініями — внутрішньої надійності через α Криппендорфа і зовнішньої валідності через F1 проти трьох незалежних gold standards (Human, LLM, Expanded). Зафіксована системна асиметрія між людським і машинним кодуванням засвідчує діагностичну, а не повністю класифікаційну валідність інструменту і обґрунтовує винятково макрорівневий режим інтерпретації результатів.

У сирих даних усі сім категорій демонструють статистично значущий спадний тренд за слабкої змістовної сили ефекту. Пряма жанрова стандартизація істотно послаблює тренд для п'яти категорій (OPN, CON, UA-FRE, UA-WAR, UA-IDN), вказуючи на значний внесок жанрової рекомпозиції корпусу як самостійного джерела варіації. На агрегатному рівні гіпотези H1–H3 не підтверджуються. На жанрово-стратифікованому рівні результати є складнішими: H1s частково підтримується локальним піком UA-IDN у реалістичній прозі (блок А) пост-помаранчевого періоду 2004–2013 рр.; H2s частково підтримується стійким зростанням ENH у жанровій белетристиці (блок В); H3s не підтверджується в первісно очікуваному вигляді для блоку А, оскільки максимум TRA у реалістичній прозі припадає на 2014–2021 рр., а не на 2022–2025 рр., проте має часткову підтримку в блоках В і D, де самотрансцендентні маркери зростають у довшій перспективі. Різні сегменти літературного поля мають істотно відмінні ціннісні траєкторії, які агрегатний аналіз нівелює. Тим самим жанрова стратифікація постає не як технічна процедура, а як самостійний соціологічний інструмент аналізу культурного поля у термінах П. Бурдьє.

Один із головних соціологічних результатів дослідження полягає в обґрунтуванні робочої інтерпретації розходження між корпусними та опитувальними показниками як можливої дисоціації режимів репрезентації

цінностей. Тріангуляція корпусних результатів з опитувальними оптиками виявила три патерни: синхронну згоду для ENH (CONVERGENT), синхронну стабільність для TRA (FLAT) і стійку дивергенцію між декларативним зростанням і зниженням лексичних маркерів для п'яти з семи категорій (DIVERGENT). Найвиразніший випадок — UA-IDN: WVS і EVS у 2020 р. дають для України практично ідентичне значення індикатора G006 (72,7 vs 72,8 пункту), тоді як стандартизована активність корпусу падає з 11,1% до 7,9%. Цей патерн не дає підстав автоматично говорити про «ціннісну кризу» або послаблення відповідних орієнтацій; натомість він дозволяє припустити зміну способу їхнього текстового вираження: частина орієнтацій, що в ранніх жанрових сегментах прямо лексично маркувалася, у пізніших текстах може переходити у фонові, сюжетно-наративні та дискурсивно опосередковані форми. У межах культурно-соціологічної традиції такі репрезентації можна розглядати як один із механізмів соціокультурної фіксації та циркуляції ціннісних смислів у публічному художньому дискурсі; тому реконструкція їхньої динаміки має самостійне соціологічне значення, але не є прямим вимірюванням ціннісних орієнтацій суспільства.

Наукова новизна полягає у розробленні інтегрованого текстоцентричного підходу, обґрунтуванні жанрово-чутливої процедури інтерпретації трендів, розвитку гібридної стратегії валідації автоматизованого аксіологічного кодування та попередній емпіричній концептуалізації дисоціації режимів репрезентації як перспективної аналітичної проблеми соціології культури. Перспективи подальшого дослідження визначаються трьома напрямками: структурна ревізія словника UA-IDN; повна статистична перевірка жанрово-стратифікованих гіпотез H1s–H3s з охопленням усіх жанрових блоків; розширення інструментарію на маркери нелексикалізованих режимів репрезентації — сюжетні структури морального вибору, фокалізацію, інтертекстуальні відсилання та символічні конфігурації.

РЕКОМЕНДАЦІЇ

На основі отриманих теоретичних, методологічних та емпіричних результатів дослідження сформульовано рекомендації, адресовані академічній спільноті, освітнім інституціям та аналітичним центрам, що працюють із культурною політикою й моніторингом ціннісних трансформацій.

1. Для подальших соціологічних і міждисциплінарних досліджень доцільно провести структурну ревізію словника UA-IDN із залученням сучасного шару української ідентичнісної лексики 2014–2025 рр.: маркерів практик спротиву, громадянської суб'єктності, дистанціювання від російського культурно-політичного простору, мовчазної нормалізації українськості. Аналогічна, але менш масштабна ревізія потрібна для словників UA-FRE і UA-WAR.
2. Для підвищення чутливості методу рекомендовано доповнити автоматизоване лексичне кодування контекстними правилами розпізнавання нелексикалізованих режимів репрезентації цінностей: сюжетних структур морального вибору, фокалізації, інтертекстуальних відсилань, символічних структур. Це дозволить адекватніше реагувати на виявлений у дослідженні зсув режиму репрезентації, який не редукується до лексично-словникового рівня.
3. Сформований корпус, кодбук, категоріальну систему та процедури валідації доцільно використовувати в подальших дослідженнях українського художнього дискурсу, а також у навчальних курсах із соціології культури, методології соціологічного дослідження, аналізу соціальних даних, контент-аналізу та цифрових гуманітарних студій. Гібридну валідаційну стратегію (Human-gold + LLM-gold + Expanded-gold) доцільно розглядати як перспективну методичну схему для україномовних автоматизованих процедур аналізу латентного смислу художнього тексту,

особливо на етапі пілотної перевірки словників, категорій і правил кодування. Для забезпечення відтворюваності результатів рекомендовано зберігати корпус, кодбук, словники маркерів, скрипти обробки, таблиці валідації та підсумкові агреговані датасети у версіонованому вигляді. Окремо доцільно документувати зміни словників між ітераціями, параметри лематизації, правила виключення або включення текстів, а також межі використання інструмента на макро- і мікрорівні. Це дасть змогу не лише повторити розрахунки, а й коректно порівнювати результати майбутніх досліджень із поточною версією корпусного аналізу

4. Для аналітичних центрів і дослідницьких інституцій, що моніторять культурні трансформації, рекомендовано використовувати поєднання корпусних і опитувальних оптик як комплементарних, а не взаємозамінних джерел: розходження між ними не слід автоматично трактувати як артефакт помилкового вимірювання; за умови методологічного контролю воно може розглядатися як самостійний діагностичний показник зміни режиму репрезентації цінностей.
5. Подальші дослідження мають розширити жанрово-стратифікований аналіз із повною статистичною перевіркою гіпотез $H1s-H3s$ у кожному жанровому блоці, систематизувати якісний шар (не менше 30–50 фрагментів для повного покриття матриці період $\times$ жанр) та у довшій перспективі — за окремим дослідницьким дизайном — залучити порівняльний матеріал інших постсоціалістичних літератур для перевірки того, чи є виявлений патерн дисоціації режимів репрезентації специфічно українським або ширшим регіональним явищем.

Реалізація запропонованих рекомендацій сприятиме поглибленню соціологічного розуміння динаміки ціннісних репрезентацій в українському художньому дискурсі та розвитку україномовного інструментарію автоматизованого аксіологічного аналізу.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Бакіров, В. С., & Сокур'янська, Л. Г. (2023). Ціннісні зрушення у повоєнній Україні: напрямки, чинники, суб'єкти. Вісник Харківського національного університету ім. В. Н. Каразіна. Серія «Соціологічні дослідження сучасного суспільства», (51), 29–37. <https://doi.org/10.26565/2227-6521-2023-51-03>

Білик, І. І. (1993). Не дратуйте грифонів: Роман. Заповіт.

Бовсунівська, Т. В. (2009). Теорія літературних жанрів: жанрова парадигма сучасного зарубіжного роману: Підручник. ВПЦ «Київський університет».

Бовсунівська, Т. В. (2010). Когнітивна жанрологія і поетика: Монографія. Київський університет.

Головаха, Є. І. (2016). Українське суспільство: шляхи трансформації. *Ukrainian Sociological Journal*, 19(1–2), 26–30.

Головаха, Є. І., Костенко, Н. В., & Макєєв, С. О. (Ред.). (2014). Суспільство без довіри. ІС НАН України.

Головаха, Є. І., & Макєєв, С. О. (Ред.). (2022). Українське суспільство в умовах війни. 2022: Колективна монографія. Інститут соціології НАН України.

Головаха, Є., & Макєєв, С. (Ред.). (2023). Українське суспільство в умовах війни. Рік 2023: Колективна монографія. Інститут соціології НАН України.

Головаха, Є. І., & Паніна, Н. В. (1994). Соціальне безумство: Історія, теорія та сучасна практика. Абрис.

Головаха, Є. І., & Паніна, Н. В. (2001). Пострадянська деінституціоналізація та становлення нових соціальних інститутів в Україні. *Соціологія: теорія, методи, маркетинг*, (4), 5–22.

Гундорова, Т. (2012). Транзитна культура. Симптоми постколоніальної травми: Есеї. Грані-Т.

Качак, Т. Б. (2017–2018). Особливості жанрової системи сучасної української прози для дітей та юнацтва. Прикарпатський вісник НТШ. Слово, 4(40)–3(47), 229–240. http://nbuv.gov.ua/UJRN/Pvntsh_sl_2017-2018_4-3_24

Київський міжнародний інститут соціології. (2024, 10 грудня). Сприйняття належності до української нації. <https://www.kiis.com.ua/?cat=reports&id=1458&lang=ukr&page=1>

Книжкова палата України імені Івана Федорова. (n.d.). Літописи. <https://ukrbook.net/litopysy.html>

Комітет з Національної премії України імені Тараса Шевченка. (n.d.). Офіційний сайт. <https://knpu.gov.ua>

Конкурс імені Григора Тютюнника. (n.d.). [Facebook page]. Facebook. <https://www.facebook.com/HryhirTiutiunnykContest/> (Архівна копія: <https://web.archive.org/web/20240615070259/https://www.facebook.com/HryhirTiutiunnykContest/>)

Коронація слова. (n.d.). Міжнародний літературний конкурс. <https://koronatsiya.com>

Костенко, Л. В. (2010). Записки українського самашедшого. А-БА-БА-ГА-ЛА-МА-ГА.

Криштопа, О., & Каліновська, О. (2016). Герої (не)війни. Про мертвих, живих і ненароджених. Книжковий Клуб «Клуб Сімейного Дозвілля».

Лис, В. С. (2019). Обітниця: Роман. Книжковий Клуб «Клуб Сімейного Дозвілля».

Лис, В. С. (2022). Вифлеєм: Роман-триптих (Серія «Великий роман»). Фоліо.

Маїлян, А. С. (2025). Комплексне застосування кількісних і якісних методів аналізу тексту в ЗМІ для визначення можливості впливу на настрої, емоційний стан та політичну спрямованість читачів. У *Україна у просторі соціальних комунікацій: Суспільство, медіа, рефлексії війни: Матеріали конференції* (С. 672–675). Державний торговельно-економічний університет. <https://doi.org/10.31617/k.knute.2025-02-27-28>

Маїлян, А. С. (2026). Методи аналізу ціннісних систем у художньому дискурсі: текстоцентричний підхід. *Вісник науки та освіти. Серія: Філологія*. [https://doi.org/10.52058/2786-6165-2026-4\(46\)-1444-1491](https://doi.org/10.52058/2786-6165-2026-4(46)-1444-1491)

Новік, А. (2022). *Медальйон трьох змій: Роман* (М. Колосова, Іл.). Nebo BookLab Publishing.

Паніна, Н. В. (2001). Молодь України: структура цінностей, соціальне самопочуття та морально-психологічний стан в умовах тотальної аномії. *Соціологія: теорія, методи, маркетинг*, (1), 5–26.

Разумков Центр. (2026). Оцінка змін, що відбулися в українському суспільстві за час повномасштабної війни (листопад 2025 р.). <https://razumkov.org.ua/napriamky/sotsiologichni-doslidzhennia/otsinka-zmin-shcho-vidbulysia-v-ukrainskomu-suspilstvi-za-chas-povnomasshtabnoi-viiny-lystopad-2025r>

Рафеєнко, В. (2023). *Петрикор — запах землі після дощу: Роман*. Видавництво Анетти Антоненко.

Соціологічна група «Рейтинг», & Gallup International. (2023). End of Year 2023: Всеукраїнське опитування. <https://ratinggroup.ua/research/ukraine/riven-demokratiyi-doslidzhennya-rejtingu-z-gallup-international.html>

Український центр європейської політики, & Info Sapiens. (2020). Світове дослідження цінностей 2020 в Україні: Аналітичний звіт. https://ucerp.org.ua/wp-content/uploads/2020/11/WVS_UA_2020_report_WEB.pdf

Фонд «Демократичні ініціативи» імені Ілька Кучеріва, & КМІС. (2025, січень). Україна єдина: національна належність, ідентичність, мова та державні атрибути: Всеукраїнське опитування. <https://dif.org.ua/article/ukraina-edina-natsionalna-nalezhnist-identichnist-mova-ta-derzhavni-atributi-vseukrainske-opituvannya>

Alexander, J. C. (2003). *The meanings of social life: A cultural sociology*. Oxford University Press.

Alexander, J. C. (2012). *Trauma: A social theory*. Polity Press.

Armitage, P. (1955). Tests for linear trends in proportions and frequencies. *Biometrics*, 11(3), 375–386.

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays* (M. Holquist, Ed.; C. Emerson & M. Holquist, Trans.). University of Texas Press.

Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics* (C. Emerson, Ed. & Trans.). University of Minnesota Press.

Bal, M. (2017). *Narratology: Introduction to the theory of narrative* (4th ed.). University of Toronto Press.

BBC News Україна. (n.d.). Книга року BBC. <https://www.bbc.com/ukrainian/topics/cqwrq3rkdrnt>

Becker, H. S. (2007). *Telling about society*. University of Chicago Press.

Berelson, B. (1952). *Content analysis in communication research*. Free Press.

Berglund, S., Ekman, J., Deegan-Krause, K., & Knutsen, T. (Eds.). (2013). *The handbook of political change in Eastern Europe* (3rd ed.). Edward Elgar.

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243–257. <https://doi.org/10.1093/lc/8.4.243>

Booth, W. C. (1983). *The rhetoric of fiction* (2nd ed.). University of Chicago Press.

Bourdieu, P. (1990). *The logic of practice* (R. Nice, Trans.). Stanford University Press.

Bourdieu, P. (1993). *The field of cultural production: Essays on art and literature*. Columbia University Press.

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). *The development and psychometric properties of LIWC-22*. University of Texas at Austin.

Brubaker, R. (1996). *Nationalism reframed: Nationhood and the national question in the new Europe*. Cambridge University Press.

Chaplynskyi, D. (2021). The infrastructure of Ukrainian language processing: Lang-uk initiative. In *Proceedings of the 1st International Workshop on Computational Linguistics for Ukrainian (UNLP 2021)*. <https://lang.org.ua>

Chatman, S. (1978). *Story and discourse: Narrative structure in fiction and film*. Cornell University Press.

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics*, 10(4), 417–451.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Coser, L. A. (1963). *Sociology through literature: An introductory reader*. Prentice-Hall.

Cramér, H. (1946). *Mathematical methods of statistics*. Princeton University Press.

Das Gupta, P. (1993). Standardization and decomposition of rates: A user's manual (Current Population Reports, P23-186). U.S. Census Bureau.

de Beaugrande, R., & Dressler, W. (1981). Introduction to text linguistics. Longman.

Durkheim, E. (1982). The rules of sociological method (W. D. Halls, Trans.). Free Press. (Original work published 1895)

Durkheim, E. (1984). The division of labour in society (W. D. Halls, Trans.). Macmillan. (Original work published 1893)

Durkheim, E. (1995). The elementary forms of religious life (K. E. Fields, Trans.). Free Press. (Original work published 1912)

English, J. F. (2005). The economy of prestige: Prizes, awards, and the circulation of cultural value. Harvard University Press.

European Values Study. (n.d.). <https://europeanvaluesstudy.eu/>

European Values Study. (2022). EVS Trend File 1981–2017 (ZA7503 Data file Version 3.0.0) [Data set]. GESIS Data Archive. <https://doi.org/10.4232/1.14021>

Fairclough, N. (1992). Discourse and social change. Polity Press.

Fairclough, N. (2003). Analysing discourse: Textual analysis for social research. Routledge.

Fileva, E. (2023). Corpus building for literary history. In C. Schöch, J. Dudar, & E. Fileva (Eds.), Survey of methods in computational literary studies. CLS INFRA. <https://doi.org/10.5281/zenodo.7892112>

Fischer, R., Karl, J. A., Luczak-Roesch, M., Fetvadjiev, V. H., & Grener, A. (2022). Opportunities and challenges of extracting values in autobiographical narratives. *Frontiers in Psychology*, 13, Article 886455. <https://doi.org/10.3389/fpsyg.2022.886455>

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Addison-Wesley.

Foucault, M. (1972). *The archaeology of knowledge* (A. M. Sheridan Smith, Trans.). Routledge.

Genette, G. (1997a). *Palimpsests: Literature in the second degree* (Ch. Newman & C. Doubinsky, Trans.). University of Nebraska Press.

Genette, G. (1997b). *Paratexts: Thresholds of interpretation* (J. E. Lewin, Trans.; R. Macksey, Foreword). Cambridge University Press.

Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30), Article e2305016120. <https://doi.org/10.1073/pnas.2305016120>

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton University Press.

Griswold, W. (2013). *Cultures and societies in a changing world* (4th ed.). SAGE/Pine Forge Press.

Grushetsky, A., & Paniotto, V. (2025). *War and the transformation of Ukrainian society (2022–23): Empirical evidence*. ibidem-Verlag.

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., & Puranen, B. (Eds.). (2022a). *World Values Survey: All rounds — Country-pooled datafile* (Version: WVS_TimeSeries_1981_2022_v5.0) [Data set]. JD Systems Institute & WVSA Secretariat. <https://doi.org/10.14281/18241.24>

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., & Puranen, B. (Eds.). (2022b). *World Values Survey Trend File (1981–2022): Cross-national data-set* (Version 4.0.0) [Data set]. JD Systems Institute & WVSA Secretariat. <https://doi.org/10.14281/18241.27>

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89. <https://doi.org/10.1080/19312450709336664>

Hitlin, S., & Piliavin, J. A. (2004). Values: Reviving a dormant concept. *Annual Review of Sociology*, 30, 359–393. <https://doi.org/10.1146/annurev.soc.30.012703.110640>

Howard, M. M. (2002). The weakness of postcommunist civil society. *Journal of Democracy*, 13(1), 157–169.

Howard, M. M. (2003). *The weakness of civil society in post-communist Europe*. Cambridge University Press.

Hunston, S., & Thompson, G. (Eds.). (2000). *Evaluation in text: Authorial stance and the construction of discourse*. Oxford University Press.

Inglehart, R. (1977). *The silent revolution: Changing values and political styles among Western publics*. Princeton University Press.

Inglehart, R. (1997). *Modernization and postmodernization: Cultural, economic, and political change in 43 societies*. Princeton University Press.

Inglehart, R., & Welzel, C. (2005). *Modernization, cultural change, and democracy: The human development sequence*. Cambridge University Press.

Kitagawa, E. M. (1964). Standardized comparisons in population research. *Demography*, 1(1), 296–315.

Kluckhohn, C. (1951). Values and value-orientations in the theory of action: An exploration in definition and classification. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action* (pp. 388–433). Harvard University Press.

Kluckhohn, F. R., & Strodtbeck, F. L. (1961). *Variations in value orientations*. Row, Peterson.

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE.

Kristeva, J. (1980). *Desire in language: A semiotic approach to literature and art* (L. S. Roudiez, Ed.; T. Gora, A. Jardine, & L. S. Roudiez, Trans.). Columbia University Press.

Kulyk, V. (2016). National identity in Ukraine: Impact of Euromaidan and the war. *Europe-Asia Studies*, 68(4), 588–608. <https://doi.org/10.1080/09668136.2016.1174980>

Kuzio, T. (2010). Nationalism, identity and civil society in Ukraine: Understanding the Orange Revolution. *Communist and Post-Communist Studies*, 43(3), 285–296. <https://doi.org/10.1016/j.postcomstud.2010.07.001>

Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). University of Washington Press.

lang-uk. (n.d.-a). *Tone-Dict-Ukrainian: Словник тональності української мови* [Computer software repository]. GitHub. <https://github.com/lang-uk/tone-dict-uk>

lang-uk. (n.d.-b). *word2vec-моделі Уберкорпусу: Навчені моделі векторних представлень* [Data set]. <https://lang.org.ua/static/downloads/models/ubercorpus.lowercased.lemmatized.word2vec.300d.bz2>

Lin, M., Lucas, H. C., & Shmueli, G. (2013). Too big to fail: Large samples and the p-value problem. *Information Systems Research*, 24(4), 906–917.

Liu, B. (2012). Sentiment analysis and opinion mining. Morgan & Claypool.
<https://doi.org/10.2200/S00416ED1V01Y201204HLT016>

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587–604. <https://doi.org/10.1111/j.1468-2958.2002.tb00826.x>

Longo, M. (2016). Fiction and social reality: Literature and narrative as sociological resources. Routledge. <https://doi.org/10.4324/9781315582269>

Lotman, Yu. M. (1977). The structure of the artistic text. University of Michigan, Department of Slavic Languages and Literatures.

Lotman, Yu. M. (1990). Universe of the mind: A semiotic theory of culture (A. Shukman, Trans.). I.B. Tauris.

Mailian, A., & Korzhov, H. (2026, April 16). Analysis of value narratives in fiction as a factor of social influence in the digital media environment. In *Фундаментальні та прикладні проблеми суспільства: історія, сьогодення, майбутнє: Матеріали Міжнародної науково-практичної конференції* (С. 57–60). Державний торговельно-економічний університет.
<https://doi.org/10.31617/k.knute.2026-04-16>

Maio, G. R. (2016). The psychology of human values. Routledge.

Mannheim, K. (1936). Ideology and utopia: An introduction to the sociology of knowledge. Routledge & Kegan Paul.

Mannheim, K. (1952). The problem of generations. In P. Kecskemeti (Ed.), *Essays on the sociology of knowledge* (pp. 276–322). Routledge & Kegan Paul.

Maslow, A. H. (1954). Motivation and personality. Harper & Row.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26, 3111–3119.

Mishler, W., & Rose, R. (2001). What are the origins of political trust? Testing institutional and cultural theories in post-communist societies. *Comparative Political Studies*, 34(1), 30–62.

Moretti, F. (2005). *Graphs, maps, trees: Abstract models for a literary history*. Verso.

Nelson, L. K. (2020). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, 49(1), 3–42. <https://doi.org/10.1177/0049124117729703>

Neuendorf, K. A. (2017). *The content analysis guidebook* (2nd ed.). SAGE.

Offe, C. (1991). Capitalism by democratic design? Democratic theory facing the triple transition in East Central Europe. *Social Research*, 58(4), 865–892.

Ogburn, W. F. (1922). *Social change with respect to culture and original nature*. B. W. Huebsch.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135.

Parsons, T. (1951). *The social system*. Free Press.

Parsons, T., & Shils, E. A. (Eds.). (1951a). *Toward a general theory of action*. Harvard University Press.

Parsons, T., & Shils, E. A. (1951b). Values, motives, and systems of action. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action* (pp. 47–275). Harvard University Press.

PEN Ukraine, & Сенсор. (n.d.). 100 книжок, які допоможуть зрозуміти Україну. <https://pen.org.ua/100-knyzhok-yaki-dopomozhut-zrozumity-ukrayinu>

Pop-Eleches, G., & Tucker, J. A. (2017). *Communism's shadow: Historical legacies and contemporary political attitudes*. Princeton University Press.

Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020). Stanza: A Python natural language processing toolkit for many human languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 101–108). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-demos.14>

Razumkov Centre. (2024a). Ідентичність громадян України: тенденції змін (червень 2024 р.) [Соціологічне дослідження]. <https://razumkov.org.ua/napriamky/sotsiologichni-doslidzhennia/identychnist-gromadian-ukrainy-tendentsii-zmin-cherven-2024r>

Razumkov Centre. (2024b). Рівень релігійності, довіра до Церкви, конфесійний розподіл та міжцерковні відносини в українському суспільстві (жовтень 2024 р.). <https://razumkov.org.ua/napriamky/sotsiologichni-doslidzhennia/riven-religiinosti-dovira-do-tserkvy-konfesiinyi-rozpodil-ta-mizhtserkovni-vidnosyny-v-ukrainskomu-suspilstvi-zhovten-2024r>

Razumkov Centre. (2025). Оцінка ситуації в країні, довіра до соціальних інститутів (листопад 2025). <https://razumkov.org.ua/napriamky/sotsiologichni-doslidzhennia/otsinka-sytuatsii-v-kraini-dovira-do-sotsialnykh-instytutiv-otsinka-problem-shcho-isnuiut-v-ukrainskomu-suspilstvi-lystopad-2025r>

Reisigl, M., & Wodak, R. (2015). The discourse-historical approach (DHA). In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse studies* (3rd ed., pp. 23–61). SAGE.

Rickert, H. (1921). *System der Philosophie. Erster Teil: Allgemeine Grundlegung der Philosophie*. Mohr.

Riessman, C. K. (2008). *Narrative methods for the human sciences*. SAGE.

Rimmon-Kenan, Sh. (2002). *Narrative fiction: Contemporary poetics* (2nd ed.). Routledge.

Rohan, M. J. (2000). A rose by any name? The values construct. *Personality and Social Psychology Review*, 4(3), 255–277. https://doi.org/10.1207/S15327957PSPR0403_4

Rokeach, M. (1973). *The nature of human values*. Free Press.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm, *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). Yale University Press.

Scheler, M. (1973). *Formalism in ethics and non-formal ethics of values: A new attempt toward the foundation of an ethical personalism* (M. S. Frings & R. L. Funk, Trans.). Northwestern University Press.

Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25, 1–65. [https://doi.org/10.1016/S0065-2601\(08\)60281-6](https://doi.org/10.1016/S0065-2601(08)60281-6)

Schwartz, S. H. (2006). Basic human values: Theory, measurement, and applications. *Revue française de sociologie*, 47(4), 929–968.

Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture*, 2(1), 1–20. <https://doi.org/10.9707/2307-0919.1116>

Schwartz, S. H. (2021). A repository of Schwartz value scales with instructions and an introduction. *Online Readings in Psychology and Culture*, 2(2). <https://doi.org/10.9707/2307-0919.1173>

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.-E., Demirutku, K., Dirilen-Gumus, O., & Konty, M. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, 103(4), 663–688. <https://doi.org/10.1037/a0029393>

Shynkarov, Y., Solopova, V., & Schmitt, V. (2025). Improving sentiment analysis for Ukrainian social media code-switching data. In M. Romanyshyn (Ed.), *Proceedings of the Fourth Ukrainian Natural Language Processing Workshop (UNLP 2025)* (pp. 179–193). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.unlp-1.18>

Sztompka, P. (1993a). Civilizational incompetence: The trap of post-communist societies. *Zeitschrift für Soziologie*, 22(2), 85–95.

Sztompka, P. (1993b). *The sociology of social change*. Blackwell.

Sztompka, P. (2000). Cultural trauma: The other face of social change. *European Journal of Social Theory*, 3(4), 449–466. <https://doi.org/10.1177/136843100003004004>

Thomas, W. I., & Znaniecki, F. (1918). *The Polish peasant in Europe and America* (Vol. 1). University of Chicago Press.

Titscher, S., Meyer, M., Wodak, R., & Vetter, E. (2000). *Methods of text and discourse analysis* (B. Jenner, Trans.). SAGE.

Vachudova, M. A. (2005). *Europe undivided: Democracy, leverage, and integration after communism*. Oxford University Press.

Van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. SAGE.

Van Dijk, T. A. (2008). *Discourse and power*. Palgrave Macmillan.
<https://doi.org/10.1007/978-1-137-07299-3>

Váňa, J. (2020). Theorizing the social through literary fiction: For a new sociology of literature. *Cultural Sociology*, 14(2), 180–200.
<https://doi.org/10.1177/1749975520922469>

Weber, M. (1946a). From Max Weber: *Essays in sociology* (H. H. Gerth & C. W. Mills, Eds.). Oxford University Press.

Weber, M. (1946b). Religious rejections of the world and their directions. In H. H. Gerth & C. W. Mills (Eds.), *From Max Weber: Essays in sociology* (pp. 323–359). Oxford University Press.

Weber, M. (1978). *Economy and society: An outline of interpretive sociology* (G. Roth & C. Wittich, Eds.; 2 vols.). University of California Press.

Welzel, C. (2013). *Freedom rising: Human empowerment and the quest for emancipation*. Cambridge University Press.

Wilkerson, J., & Casas, A. (2017). Large-scale computerized text analysis in political science: Opportunities and challenges. *Annual Review of Political Science*, 20, 529–544. <https://doi.org/10.1146/annurev-polisci-052615-025542>

Windelband, W. (1903). *Präludien: Aufsätze und Reden zur Philosophie und ihrer Geschichte* (2nd ed.). Mohr.

World Values Survey Association. (n.d.-a). *World Values Survey — Wave 3* (Ukraine, 1996). <https://www.worldvaluessurvey.org/WVSDocumentationWV3.jsp>

World Values Survey Association. (n.d.-b). World Values Survey — Wave 5 (Ukraine, 2006). <https://www.worldvaluessurvey.org/WVSDocumentationWV5.jsp>

World Values Survey Association. (n.d.-c). World Values Survey — Wave 6 (Ukraine, 2011). <https://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp>

World Values Survey Association. (n.d.-d). World Values Survey official website. <https://www.worldvaluessurvey.org/>

Wuthnow, R. (1987). *Meaning and moral order: Explorations in cultural analysis*. University of California Press.

Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., & Yang, D. (2024). Can large language models transform computational social science? *Computational Linguistics*, 50(1), 237–291. https://doi.org/10.1162/coli_a_00502

ДОДАТКИ

Додаток А. Концепції цінностей в соціології

Таблиця А.1.

Ключові соціологічні концепції цінностей та їх операціоналізація в дослідженні художнього дискурсу

Автор / концепція	Рівень аналізу	Ключова аналітична одиниця	Застосування в цьому дослідженні
Дюркгейм	Суспільство / колектив	Колективні уявлення	Художній текст як матеріалізація колективних уявлень
Вебер	Дія / сфера	Ціннісна раціональність, Wertsphären	Художній дискурс як арена міжсферних ціннісних напруг
Томас і Знанецький	Соціальна група	Соціальна цінність / установка	Аналіз особистих документів як методологічний прецедент; для художнього тексту — опосередкована аналогія
Парсонс	Система / інститут	Ціннісний стандарт вибору	Цінності в художньому тексті як нормативні позиції
Рокіч	Індивід / група	Термінальні / інструментальні цінності	Виявлення розриву між декларованим і демонстрованим
Маннгейм	Покоління / соціально-історична спільність	Поколінний досвід як інтерпретаційна рамка	Хронологічні рубежі 1991, 2004, 2014, 2022 як поколіннєві зрізи
Інглхарт / Велзель	Суспільство / покоління	Вживання / самовираження	Зовнішній контекст хронологічних зрізів (WVS)
Шварц	Індивід / культура	10 базових типів, мотиваційне коло	Аналіз реконфігурацій ціннісних кластерів у корпусі
Бурдьє	Поле / практика	Хабітус, символічний капітал	Ціннісний дискурс як результат взаємодії авторської позиції, структури літературного поля і символічного капіталу

Додаток Б. Поняття та індикатори

Таблиця Б.1.

**Ключові поняття дослідження та суміжні концепти: аналітичне
розмежування і можливі індикатори в художньому тексті**

Поняття	Рівень носія або прояву	Ступінь ситуативності	Нормативний характер	Спосіб виявлення у художньому тексті
Цінність	Індивідуальний / груповий / соціокультурний	Надситуативна	Оцінний орієнтир («що є бажаним»)	Наскрізнi теми, моральні опозиції, символічні образи, оцінна лексика
Ціннісна орієнтація	Індивідуальний / груповий	Відносно стійка	Ієрархія індивідуальних або групових пріоритетів	Стійкі рішення персонажів, моральні вибори, конфлікти
Ціннісна система	Індивідуальний** / груповий / культурний / дискурсивний	Відносно стійка; може змінюватись історично або під впливом кризових подій	Структурна організація цінностей	Конфігурація й ієрархія цінностей у корпусі текстів; повторювані ціннісні опозиції
Ціннісна установка	Індивідуальний / груповий	Ситуативна	Готовність до реакції / дії	Ситуативні реакції персонажів, модальність висловлювань, оцінні предикати, індикатори тональності
Норма	Соціальний / інституційний	Відносно стійка	Правило («як належить діяти»)	Вимоги, очікування, санкції, нормативна мова оповідача

Потреба	Біологічний / психологічний / соціальний	Ситуативно актуалізована	Мотиваційний стан нестачі або розвитку («чого бракує / до чого спрямоване зростання»)	Мотиваційні рушії персонажів, зображення нестачі / задоволення
Інтерес	Індивідуальний / груповий / класовий	Ситуативно і структурно зумовлений	Практично й позиційно зумовлена спрямованість дії («що є ставкою у певному полі»)	Конфлікти між персонажами, риторика виправдання власних дій
Переконання	Індивідуальний / груповий	Відносно стійке	Когнітивне твердження («що є правдою/істиною»)	Судження персонажів і наратора, когнітивні дієслова та предикати
Ідеологія	Груповий / суспільний / дискурсивний	Відносно стійка, змінюється під впливом соціальних зрушень	Система групових переконань, цінностей і пояснювальних схем, що легітимізує або оспорує соціальний порядок	Образи «своїх» і «чужих», топоси легітимації влади, пресупозиції щодо норм, наративні схеми виправдання або осуду

**Таблицю складено автором на основі узагальнення підходів М. Рокіча, Ш. Шварца, В. Томаса і Ф. Знанецького, Т. Парсонса, А. Маслоу, П. Бурдьє, Т. ван Дейка, М. Розенберга і К. Говланда (Bourdieu, 1990, 1993; Fishbein & Ajzen, 1975; Hitlin & Piliavin, 2004; Kluckhohn & Strodtbeck, 1961; Maslow, 1954; Parsons, 1951; Rokeach, 1973; Rosenberg & Hovland, 1960; Schwartz, 1992; Thomas & Znaniecki, 1918; van Dijk, 1998).*

*** Індивідуальний рівень ціннісної системи розглядається як ієрархія особистісних пріоритетів (Rokeach, 1973; Schwartz, 1992); водночас у межах цього дослідження особливу увагу приділено груповому, культурному та дискурсивному рівням, оскільки саме вони дозволяють пов'язати ціннісні структури художнього тексту з ширшими соціокультурними процесами.*

Додаток В. Жанрова класифікація

Таблиця В.1

Жанрова класифікаційна таблиця. Блок А. Реалістична та сучасна проза

Код	Жанр	Обґрунтування включення
G1	Соціальна проза	Прямий об'єкт аксіологічного аналізу: відображення суспільних суперечностей і ціннісних конфліктів
G2	Психологічна проза	Внутрішній монолог як носій суб'єктивних ціннісних орієнтацій
G3	Автобіографічна / мемуарна проза	Декларована ціннісна система автора через нарацію реального досвіду
G4	Воєнна проза	Особливо релевантна для 2014–2025 рр.; трансформація ціннісної системи в умовах конфлікту
G5	Міська / урбаністична проза	Місто як простір ціннісної трансформації та ідентичності

Таблиця В.2

Жанрова класифікаційна таблиця. Блок В. Жанрова белетристика

Код	Жанр / Піджанр	Обґрунтування включення
G6	Детектив	Моральна опозиція «справедливість / злочин»; нормативна ціннісна рамка
G6.1	Класичний детектив	Логічне розслідування, архетипна моральна бінарність
G6.2	Кримінальний трилер	Психологічний тиск; цінність виживання і справедливості
G6.3	Поліцейський процедурал	Інституційний вимір ціннісної системи
G7	Жіноча проза	Жіноча ідентичність, самовизначення; особливо актуальна для аналізу гендерних ціннісних трансформацій. Категорія G7 «жіноча проза» використовується не як есенціалістське позначення «жіночого письма», а як умовна літературно-видавнича й дослідницька категорія для групи текстів, у яких центральними є теми жіночої суб'єктності, гендерних ролей, тілесності, сім'ї, самореалізації та соціальних очікувань.

G8	Романтична / любовна проза	Норми партнерства, сім'ї, кохання як ціннісного наративу
G9	Пригодницька проза	Цінності подолання, сміливості, солідарності

Таблиця В.3

Жанрова класифікаційна таблиця. Блок С. Фантастика і фентезі

Код	Жанр / Піджанр	Обґрунтування включення
G10	Наукова фантастика (НФ)	Технологічні утопії/дистопії як проєкція ціннісних страхів і надій суспільства
G10.1	Антиутопія	Деструктивна ціннісна система як дзеркало сучасних суспільних тривог
G11	Фентезі	Архетипна бінарна ціннісна система; апеляція до культурних першоджерел
G11.1	Епічне фентезі	Боротьба добра і зла; ієрархія цінностей
G11.2	Міське фентезі	Перетин сучасних і архаїчних ціннісних систем
G11.3	Слов'янське / українське фентезі	Апеляція до архаїчних українських цінностей; особливо релевантна для дослідження
G12	Містика / горор	Межі норми; страх і захист як ціннісні маркери

Таблиця В.4

Жанрова класифікаційна таблиця. Блок D. Дитяча та підліткова проза

Код	Жанр	Обґрунтування включення
G13	Дитяча проза (6–12 р.)	Дидактична ціннісна система; маніфестні норми поведінки
G14	Підліткова проза (12–18 р.)	Ціннісне самовизначення; перші ціннісні конфлікти

Таблиця В.5

Жанрова класифікаційна таблиця. Блок Е. Гібридні та пограничні форми

Код	Жанр	Обґрунтування включення
G15	Постмодерна проза	Деконструкція усталених ціннісних систем; іронія щодо традиційних норм
G16	Художньо-документальна проза	Ціннісна система репрезентується через літературно опрацьовані реальні події, свідчення, пам'ять і досвід; жанр включається лише за наявності художньо-нарративної організації тексту.
G17	Історична проза	Апеляція до цінностей минулого; конструювання національної ідентичності
G18	Гібридні жанри	Поєднання ознак кількох жанрів. G18 використовується лише тоді, коли твір не може бути переконливо віднесений до одного домінантного жанру або поєднує кілька жанрових кодів без очевидної ієрархії.

Додаток Г. Демонстративний приклад JSON структури твору для обробки

```
{
  "schema_version": "1.1",
  "book_id": 1,
  "meta": {
    "title": "Назва твору",
    "authors": [ { "author_id": 1, "name_display": "Ім'я Прізвище", "role": "author" } ],
    "year": 2014,
    "publisher": "Назва видавництва",
    "lang": "uk"
  },
  "analytical": { "genre_primary": "G1", "genre_block": "A", "genre_academic": "роман" },
  "corpus_tags": { "period": "2014–2021" },
  "quality": { "validation_status": "OK", "ocr_suspected": false },
  "structure": { "paragraphs_count": 1464, "words_count": 56144 },
  "sections": [
    {
      "section_id": "s001", "level": 1, "title": "Розділ 1", "paragraphs": [
        {
          "para_id": "s001_p0001",
          "text": "Перший абзац книги з нормалізованим текстом.",
          "type": "prose",
          "is_dialogue": false,
          "word_count": 7
        }
      ]
    }
  ]
}
```

Наведений приклад є демонстраційним і не відтворює повний обсяг полів реального JSON-файлу; він показує лише основні структурні блоки, релевантні для подальшої аналітичної обробки.

Додаток Д. Кількісні показники робочого корпусу української художньої літератури 1991-2025

Результати формування файлів корпусу:

- corpus_flat.csv (1 375 333 рядків)
- Parquet: corpus_flat.parquet (1 375 333 рядків)
- CSV: corpus_narrative.csv (691 966 рядків)
- CSV: corpus_dialogue.csv (683 367 рядків)
- Stats: corpus_stats.csv (664 книг)
- Genre: corpus_genre_stats.csv (102 рядків)
- Файлів оброблено: 664
- Абзаців усього: 1 375 333
- недіалоговий наративний субкорпус: 691 966 (50,3%)
- діалоги: 683 367 (49,7%)
- Книг у корпусі: 664
- Авторів у корпусі: 235
- Слів усього (сума): 44 837 606

Розподіл по підперіодах:

- 1991-2003: 234 972 абзаців, 135 книг, 9 411 907 слів
- 2004-2013: 408 358 абзаців, 224 книг, 13 076 371 слів
- 2014-2021: 514 916 абзаців, 231 книг, 16 141 096 слів
- 2022-2025: 217 087 абзаців, 74 книг, 6 208 232 слів

Розподіл по жанрових блоках:

- Блок А (Реалістична та сучасна проза): 186 книг, 344 279 абзаців
- Блок В (Жанрова белетристика): 114 книг, 255 865 абзаців
- Блок С (Фантастика і фентезі): 72 книг, 208 386 абзаців
- Блок D (Дитяча та підліткова проза): 57 книг, 106 228 абзаців
- Блок Е (Гібридні та пограничні форми): 235 книг, 460 575 абзаців

Розподіл по G-кодах (топ-10):

- G17: 141 книг
- G1: 99 книг
- G15: 87 книг
- G2: 53 книг
- G6.2: 49 книг
- G14: 37 книг
- G10: 26 книг
- G13: 20 книг
- G7: 18 книг
- G12: 16 книг

Для контролю якості підготовлених файлів розроблено скрипт валідації, що перевіряє 27 параметрів у п'яти блоках: мовна ідентичність тексту (відповідність meta.lang, наявність українськоспецифічних символів, автоматична детекція мови бібліотекою langdetect); повнота метаданих (наявність і коректність ідентифікаторів, дат, жанрових кодів, хронологічного підперіоду); структурна цілісність (унікальність ідентифікаторів абзаців, цілісність ієрархії розділів, узгодженість лічильників); якість тексту (частка коротких абзаців, OCR-артефакти, неуніфіковані типи апострофа, аномально довгі абзаци); узгодженість аналітичних полів (відповідність між задекларованими та фактичними характеристиками якості файлу). Уніфікацію апострофа виконано програмно: замінено 212 018 нестандартних символів (U+2019, U+02BC, U+2018, U+0060, U+00B4) на ASCII-апостроф (U+0027); залишковий відсоток необроблених входжень склав 0,13% і локалізований у службових метаданих файлів, що не впливає на аналіз текстового масиву.

Таблиця Д.1.

Лексичний обсяг корпусу за хронологічними зрізами

Зріз	Токенів	Унікальних лем
1991–2003	11 749 729	257 394
2004–2013	16 264 200	239 483
2014–2021	19 994 248	232 094
2022–2025	7 678 578	133 778

Таблиця Д.2.

Лексичний обсяг корпусу за жанровими блоками.

Блок	Токенів	Унікальних лем
Блок А	14 234 469	235 577
Блок В	9 373 698	142 829
Блок С	7 841 087	132 723
Блок D	3 539 175	77 308
Блок Е	20 698 326	315 993

Таблиця Д.3.

Лексичний обсяг корпусу за субкорпусами

Субкорпус	Токенів	Унікальних лем
Наратив	36 937 156	411 425
Діалог	18 749 599	254 969
Корпус увесь	55 686 755	502 102

Примітка. Розбіжність між загальною кількістю слів у Таблиці 2.3 основного тексту (44 837 606) і токенів у таблицях Д.1, Д.2, Д.3 (55 686 755) зумовлена різницею між двома способами підрахунку: (1) «слова» — змістовні словоформи з вихідного парсингу JSON-файлів корпусу, без розділових знаків і службових символів; (2) «токени» — одиниці, отримані після Stanza-токенізації, де розділові знаки виокремлюються в самостійні токени. Друга метрика використовується для обчислення лексичного покриття словника маркерів (підрозділ 2.2); перша — для опису обсягу корпусу в традиційних філологічних показниках.

Додаткова інформація, зафіксована про корпус:

Лексичне розмаїття є вищим у ранніх періодах: співвідношення унікальних лем до токенів становить приблизно 2,2% у 1991–2003 рр. проти 1,2% у 2014–2021 рр. Це може бути пов'язано як із жанровою структурою раннього корпусу, так і з більшою часткою історичної,

постмодерної та стилістично неоднорідної прози. Тому цей показник потребує інтерпретації разом із жанровим складом, а не як самостійний доказ зміни літературної мови.

Домінування блоку Е є критичною особливістю корпусу: на нього припадає близько 37% усіх токенів, а також найбільший внутрішній обсяг унікальних лем серед жанрових блоків. Оскільки частина лем повторюється між блоками, показник унікальних лем не слід трактувати як просту частку словника корпусу; він радше вказує на лексичну насиченість і тематичну різноманітність блоку. Саме тому подальший аналіз потребує контролю жанрового складу.

Асиметрія наратив/діалог — наративний субкорпус містить у 2 рази більше токенів при тих же ~50% абзаців, тобто наративні абзаци в середньому в 2 рази довші за діалогові.

Додаток Е. Приклад corpus_stats.csv

Таблиця Е.1.

Приклад таблиці corpus_stats.csv (перші 9 рядків даних)

book_id	author	title	year	period	genre_primary	genre_block	genre_academic	paragraphs_total	paragraphs_narr	paragraphs_dial	words_total	words_mean_para
1	Амеліна Вікторія	Синдром листопаду, або Номо Компатиєнс	2014	2014-2021	G1	A	роман	1464	836	628	56144	38,3
3	Амеліна Вікторія	Дім для Дома	2017	2014-2021	G17	E	роман	2587	1668	919	82364	31,8
8	Андрусяк Іван	Стефа і її Чакалка	2007	2004-2013	G13	D	повість дитяча	335	160	175	8036	24,0
11	Андрусяк Іван	Кабан дикий — хвіст великий	2010	2004-2013	G13	D	повість дитяча	433	229	204	8464	19,5
15	Андрухович Софія	Літо Мілени	2002	1991-2003	G2	A	повість	362	184	178	18378	50,8
16	Андрухович Софія	Старі люди	2003	1991-2003	G2	A	проза	517	295	222	23220	44,9
17	Андрухович Софія	Жінки їхніх чоловіків	2005	2004-2013	G7	B	проза	803	789	14	33155	41,3
18	Андрухович Софія	Сьомга	2007	2004-2013	G15	E	роман	1900	1900	0	92164	48,5
19	Андрухович Софія	Фелікс Австрія	2014	2014-2021	G17	E	роман	1628	1628	0	63399	38,9

Додаток Ж. Структура полів corpus_flat.csv

Таблиця Ж.1.

Структура основних полів таблиці corpus_flat.csv

Поле	Тип	Опис	Приклад
book_id	integer	Унікальний ідентифікатор твору	1
author_id	integer	Унікальний ідентифікатор автора	12
para_id	string	Локальний ідентифікатор абзацу в межах окремого твору; для міжкнижкового аналізу унікальним ключем є комбінація book_id + para_id	s001_p0001
title	string	Назва твору	Дім для Дома
author	string	Автор	Вікторія Амеліна
year	integer	Рік видання	2017
period	string	Хронологічний зріз	2014–2021
genre_primary	string	Основний жанровий код	G17
genre_block	string	Жанровий блок	E
is_dialogue	boolean	Ознака діалогового фрагмента	true/false
word_count	integer	Кількість слів в абзаці	38
text	string	Текст абзацу	«... текст...»

Службові поля, що з'являються не у вихідному corpus_flat.csv, а у файлі результатів автоматизованого кодування після прогону словника, описано в Додатку И, розділ 6.4.

Додаток II. Кодбук дослідження

Кодбук дослідження

(операційні визначення категорій, словники маркерів, правила розмежування, процедура ручної розмітки)

1. Загальні положення

Кодбук фіксує операційні визначення сімох ціннісних категорій, словники маркерів, правила розмежування суміжних категорій, каталог граничних випадків і процедуру ручного кодування контрольної вибірки V1. Документ виконує дві функції: інструктивну (основа для навчання кодувальників-валідаторів) і специфікаційну (технічна основа для автоматизованого прогону по корпусу). Архітектурний опис категоріальної системи і обґрунтування методологічних рішень наведені у підрозділі 2.2 основного тексту; у цьому Додатку зосереджено реалізаційні деталі. Кодбук застосовується з урахуванням чотирьох методологічних меж автоматизованого підходу, описаних у підрозділі 2.2.

2. Сім ціннісних категорій

Категоріальна сітка містить сім бінарних категорій у двох групах. Група А (адаптована теорія Шварца): OPN, CON, TRA, ENH. Група В (специфіка українського контексту): UA-IDN, UA-WAR, UA-FRE.

2.1. OPN — Відкритість до змін

Агрегує цінності Самостійності, Стимуляції і Гедонізму. Стосується індивідуальної свободи вибору, новизни, ризику, творчості, насолоди, кохання як індивідуального переживання. Ключова ознака — індивідуальний суб'єкт.

Маркери: свобода (особиста), вибір, обирати, мрія, ризик, пригода, новизна, насолода, кохання, втекти, звільнитися, творити, експериментувати, самостійність, автономія.

«ТАК»: «Вона мала право на свої помилки і своє щастя».

«ТАК»: «Він знав одне — треба тікати з цього міста, де його змушують жити за чужими правилами».

«НІ»: «Народ виборів незалежність ціною крові» (свобода колективна >> UA-FRE).

«НІ»: «У нього було вільне місце у поїзді» (технічне вживання).

2.2. CON — *Збереження, родина, традиції*

Агрегує цінності Безпеки, Конформності і Традиції. Стосується родини, дому, традицій, дотримання звичаїв, безпеки, стабільності, поваги до старших і покори громадським нормам.

Маркери: родина, сім'я, мати, батько, дім, домівка, хата, рідня, обов'язок, традиція, звичай, обряд, віра, церква, ікона, рушник, весілля, поминки, спадщина, безпека, лад, спокій, повага, покора.

«ТАК»: «Мати поправила рушник на іконі і сіла за стіл. Так робили в цій хаті завжди».

«НІ»: «Він зайшов додому, повісив куртку» — «дім» у побутовому, не ціннісному значенні.

2.3. TRA — *Турбота про інших, доброта, справедливість*

Агрегує цінності Доброзичливості і Універсалізму. Стосується допомоги іншим, співчуття, милосердя, дружби, справедливості, самопожертви заради ближнього або принципу. Любов як турбота — сюди; романтична любов — до OPN.

Маркери: допомога, врятувати, віддати, ділитися, дружба, друг, співчуття, милосердя, доброта, прощення, жертвувати, відданість, справедливість, чесність, людяність, толерантність.

«ТАК»: «Бабуся віддала їй останній шматок хліба, а сама лягла спати голодна».

«НІ»: «Вони палко цілувалися всю ніч» (романтика >> OPN).

2.4. ENH — *Успіх, влада, престиж*

Агрегує цінності Досягнення і Влади. Стосується індивідуального успіху, перемоги над суперником, статусу, грошей, визнання. Колективні досягнення нації — до UA-FRE/UA-WAR.

Маркери: успіх, перемога, виграти, тріумф, амбіція, кар'єра, гроші, багатство, влада, керувати, наказати, престиж, статус, визнання, найкращий, домінування, конкуренція, суперник.

«ТАК»: «Він добився посади директора, купив дороге авто і відкрив рахунок у швейцарському банку».

«НІ»: «Україна виграла війну за свою свободу» (колективна перемога >> UA-FRE).

2.5. UA-IDN — *Українська ідентичність, культурна самобутність*

Стосується самовизначення як українця, цінності української мови, культури, історії, символіки. Герой ототожнює себе з нацією, повертається до коренів, шанує українських історичних і культурних діячів.

Маркери: українець, україномовний, національний, козак, гетьман, Січ, Україна, нація, народ, мова, рідна земля, Шевченко, Леся (Українка), Франко, вишиванка, рушник (з культурним контекстом), писанка, тризуб, батьківщина, ідентичність, самобутність.

«ТАК»: «Він повернувся до української мови, бо це мова його прадідів».

«ТАК»: «Її вишиванка з маминими стібками — це більше, ніж одяг».

«НІ»: «Він жив у Києві, на проспекті Перемоги» (географічна назва).

«НІ»: «Національний банк підвищив ставку» (назва інституту).

2.6. UA-WAR — *Війна, репресії, історичні катастрофи*

Охоплює досвід війни, репресій, Голодомору, Чорнобиля, переміщення як основу колективного «ми». Включає Другу світову, АТО/ООС, повномасштабне вторгнення 2022 р., політичні репресії, голодомори.

Маркери: війна, фронт, окоп, обстріл, ракета, тривога, окупація, окупант, ворог, біженець, переселенець, евакуація, загиблий, поранений, втрата, Голодомор, голод, репресія, ГУЛАГ, табір, депортація, Чорнобиль, Бахмут, Маріуполь, травма, пам'ять.

«ТАК»: «Бабуся переписала рецепт хліба, який вона їла у тридцять третьому» (Голодомор).

«НІ»: «У них почалася холодна війна на роботі» (метафорична війна).

2.7. UA-FRE — *Свобода нації, спротив, революція*

Охоплює національну (а не індивідуальну) свободу, спротив окупації, революційні події, боротьбу за державність. Ключова ознака — колективний суб'єкт: «ми, українці», «наш народ», «наша боротьба».

Маркери: визволення, гідність, Майдан, революція гідності, спротив, опір, повстання, суверенітет, незалежність (нації), барикада, віче, європейський вибір, не здамося, ми вистояли, голос народу, демократія (політична), помаранчева революція.

«ТАК»: «Ми стояли на Майдані три місяці і вистояли».

«ТАК»: «Український народ не здається перед загарбником».

«НІ»: «Він хотів свободи — щоб ніхто не вказував йому, як жити» (індивідуальна >> OPN).

«НІ»: «Майдан Незалежності — наша зустріч у середу» (адреса).

3. Правила кодування

Одиниця запису — абзац типу prose з мінімальною довжиною у п'ять слів. Одиниця кодування — окремий текстовий маркер у межах абзацу. Для кожного абзацу за кожною з семи категорій незалежно ухвалюється бінарне рішення: 1 (категорія присутня) або 0 (відсутня).

Категорії кодуються незалежно одна від одної: один абзац може отримати від 0 до 7 кодів. Множинне кодування є нормою: 51,4% абзаців корпусу мають принаймні одну активну категорію, середня кількість активних категорій на абзац — близько 1,5. Типовий приклад множинного кодування — абзац про матір, що готує валізу для евакуації під обстрілом, отримує TRA, CON і UA-WAR одночасно.

Кодується тільки те, що видно у тексті абзацу (з урахуванням контексту попереднього і наступного абзацу). Загальна оцінка стилю автора, літературна якість, латентний смисл або прихована мотивація — не кодуються. Носій ціннісного судження (наратор, протагоніст, антагоніст, голос у пам'яті, голос у мріях) не впливає на бінарне рішення: якщо маркер у тексті присутній, категорія кодується як 1 незалежно від оцінки автора. Згадки минулого, мрії, припущення також активують категорію за наявністю маркера.

4. Правила розмежування суміжних категорій

Низка категорій має лексично перетинні маркери. Розмежування здійснюється через рівень суб'єкта і контекст. Якщо обидві категорії одночасно відповідають змісту абзацу, кодуються обидві.

4.1. «Свобода»: OPN vs UA-FRE

Слова «свобода», «вільність», «незалежність», «воля» лексично належать до обох категорій. Розмежування — через рівень суб'єкта:

- Свобода окремої людини (я, він, вона, себе) >> OPN.
- Свобода народу/нації/країни (ми, народ, нація, українці) >> UA-FRE.

- Свобода особи в умовах національної несвободи >> обидві категорії.

У автоматичному прогоні розмежування реалізовано через службове поле `subject_level_auto` (див. розділ 6.3).

Деталізовані правила автоматичного присвоєння значень `subject_level_auto`. Поле `subject_level_auto` обчислюється автоматично для кожного абзацу, у якому виявлено лему категорії OPN або UA-FRE, і набуває одного з п'яти значень:

- **individual** — у абзаці наявні маркери індивідуального суб'єкта («я», «він», «вона», «себе», конкретне ім'я персонажа) і відсутні маркери колективного суб'єкта;
- **collective** — наявні маркери колективного суб'єкта («ми», «народ», «нація», «українець», «спільнота», «суспільство»);
- **institutional** — суб'єктом висловлювання виступає інституція (держава, армія, церква);
- **mixed** — одночасно присутні маркери індивідуального і колективного суб'єкта;
- **unclear** — жодний тип маркерів суб'єкта не зафіксовано однозначно у вікні ± 5 токенів від леми.

Правила застосування за значеннями `subject_level_auto`:

- `individual` + лема «свобода» / «вільність» / «незалежність» / «воля» >> OPN з повною вагою;
- `collective` + лема «свобода» / «незалежність» / «воля» >> UA-FRE з повною вагою;
- `institutional` + лема «свобода» >> не активує OPN (зазвичай інституційний контекст і політична риторика); кодується лише за наявності інших категорій;
- `mixed` >> подвійне кодування (OPN + UA-FRE) із зниженою вагою кожної категорії (0,5 від базової strong/standard ваги);
- `unclear` >> позначка multivalent: кейс передається у вибірку для якісної інтерпретації (Розділ 4 основного тексту, підрозділ 4.4).

Правила застосовуються після завершення лематизації і визначення `subject_level_auto`; повторні правки `subject_level` вручну не виконуються (рішення приймається автоматизованим скриптом для забезпечення відтворюваності між прогонами і незалежності від суб'єктивної інтерпретації кодувальника).

4.2. «Перемога»: ENH vs UA-FRE/UA-WAR

- Особиста перемога (виграти турнір, отримати підвищення) >> ENH.

- Колективна перемога нації (перемога народу, над загарбником) >> UA-FRE або UA-WAR.

4.3. «Традиція»: CON vs UA-IDN

- Загальна побутово-родинна або церковна традиція без національного контексту >> CON.
- Українська національна традиція з культурним контекстом (вишиванка-символ, козацький звичай, Шевченків заповіт) >> UA-IDN.
- Українська родинна традиція з культурним навантаженням >> обидві.

4.4. «Дім» і «допомога»

- Дім як цінність роду і спадкоємності >> CON. Дім як географічна локація («зайшов додому») >> не кодується. Втрата дому в умовах війни >> UA-WAR (часто разом з CON як крах цінності).
- Самопожертва і безкорислива допомога >> TRA. Простий обмін або професійна послуга без морального акценту >> не кодується.

5. Процедура ручного кодування для валідації V1

5.1. Інструмент і режим розмітки

Ручна розмітка контрольної вибірки V1 (100 абзаців за схемою 4 періоди × 5 жанрових блоків × 5 абзаців на комірку) виконується у спеціально розробленому веб-додатку `coding_app_v5d.html`. Додаток працює локально у будь-якому сучасному браузері, не потребує підключення до інтернету і не передає даних на сервери; дані зберігаються у `LocalStorage` браузера до експорту в CSV. Це рішення замінює раніше передбачений сценарій роботи з CSV-таблицями у Microsoft Excel або LibreOffice Calc, відхилений через ризики помилок ручного заповнення і втрати кодування UTF-8.

Розмітка виконується у режимі `blind`: кодувальник бачить тільки текст абзацу та сусідній контекст (попередній і наступний абзаци), але не бачить періоду, року, жанрового блоку, автора, назви твору або результатів автоматичного кодування. Це усуває упередження, пов'язане з очікуваннями щодо динаміки. Технічно сліпий режим реалізовано через два окремі файли (`coding`-файл і `metadata`-файл), які об'єднуються за `para_id` після завершення розмітки.

5.2. Інтерфейс і покрокова процедура

Інтерфейс додатку побудований навколо двоколонкової компоновки. Ліва колонка містить маркер номера абзацу, блок контексту до (згорнутий за замовчуванням), основний текст і блок контексту після. Права колонка — sticky-панель з переліком семи категорій (з короткими підказками), окремим чекбоксом «Нічого з зазначеного», полем приміток і кнопками навігації.

Для кожного абзацу кодувальник: прочитує текст; за потреби розгортає блоки контексту; для кожної з семи категорій ставить прапорець, якщо у тексті присутні відповідні маркери (за визначеннями розділу 2); якщо жодна категорія не присутня — ставить чекбокс «Нічого з зазначеного»; за потреби фіксує короткий коментар у полі notes; натискає «Далі >>» для переходу до наступного абзацу. До вже розмічених абзаців кодувальник не повертається без явної потреби.

5.3. Збереження прогресу, експорт і формат CSV

Прогрес автоматично зберігається у LocalStorage після кожного підтвердженого абзацу — закриття вкладки не призводить до втрати даних. Після проходження всіх 100 абзаців кнопка «Завершити і скачати CSV» формує файл з ім'ям V1_results_{ім'я}_{дата}_{час}.csv з BOM-розміткою UTF-8. Структура файлу:

Колонка	Зміст
para_id	Унікальний ідентифікатор абзацу (формат s{секція}_r{параграф}).
coder_name	Ім'я або кодове позначення кодувальника.
OPN...UA-FRE (7 колонок)	Бінарні значення: 1 (присутня) або 0 (відсутня).
nothing_selected	1, якщо обрано «Нічого з зазначеного».
is_skipped	1, якщо абзац пропущений.
notes	Вільний коментар кодувальника.
time_seconds, timestamp, app_version	Службові поля: час прийняття рішення, момент фіксації, версія додатку.

5.4. Організаційні умови

Орієнтовний середній час розмітки одного абзацу — 1,5–2,5 хвилини; повний обсяг 100 абзаців — 3–4 години чистої роботи. Рекомендована організація: 2–4 сесії по 45–60 хвилин з

перервами 15–30 хвилин (безперервна робота довше години знижує якість рішень). Кодувальники не обговорюють своїх рішень один з одним до завершення розмітки: координація відповідей через попередню розмову руйнує методику дослідження. У комплекті, що надається кодувальнику, є три файли: веб-додаток `coding_app_v5d.html`, функціональна інструкція `coder_instruction_V1.docx` (спрощений виклад розділів 2–4 цього кодбуку, адаптований до інтерфейсу додатку) і супровідний лист `supporting_letter.docx` з описом мети дослідження, етичних умов і технічних інструкцій.

5.5. Формування «золотого стандарту»

Фінальний «золотий стандарт» розмітки V1, проти якого порівнюється автоматичний прогон, формується через мажоритарне голосування: для кожної пари (абзац, категорія) присвоюється значення, обране щонайменше трьома з п'яти кодувальників. Випадки мінімальної більшості 3:2 фіксуються як слабкоконсесусні і позначаються окремо в діагностичних таблицях. Окрема adjudication-процедура для зміни фінального gold-рішення в межах цієї роботи не застосовується: значення `gold=1` присвоюється за правилом більшості ≥ 3 з 5 кодувальників, а статистика α Криппендорфа обчислюється на основі незалежних рішень кодувальників до будь-якого обговорення розбіжностей.

6. Технічна реалізація автоматизованого прогону

6.1. Процедура розширення стартового словника через word2vec-модель

Стартове ядро seed-маркерів кодбуку: 94 strong-леми, 92 standard-леми, 26 weak-лем, 50 n-грам — разом 262 маркери.

Технічна модель: предтренована word2vec-модель (Mikolov et al., 2013), навчена на лематизованому й нижньореєстровому Уберкорпусі (Chaplynskyi, 2021; lang-uk, n.d.-b), з 300-вимірними векторами.

Процедура. Для кожної strong- і standard-леми категорії запитуються 30 найближчих за косинусною подібністю слів. Усі кандидати агрегуються в категорію. Кожному кандидату обчислюється композитний score за чотирма компонентами:

$$\text{score} = 0,4 \cdot \text{similarity} + 0,2 \cdot \text{freq} + 0,2 \cdot \text{context} + 0,2 \cdot \text{distinct}$$

де:

- *similarity* — косинусна подібність до seed-леми;
- *freq* — нормалізована корпусна частота;

- *context* — контекстуальна узгодженість (частка спрацювань у релевантних абзацах);
- *distinct* — специфічність для категорії (мінус частка спрацювань у конкуруючих категоріях).

Ваги (0,4 + 0,2 + 0,2 + 0,2) калібровано емпірично на пілотній перевірці кандидатів категорій OPN і UA-WAR. За результатами ручного перегляду 100 пілотних кандидатів частка семантично прийнятних автоматично прийнятих лем становила близько 85%; цей показник використано як внутрішній діагностичний орієнтир, а не як формальну метрику валідності словника.

Технічні обмеження запитів. У запит до word2vec надсилаються лише однослівні strong- і standard-леми (174 з 262 маркерів стартового ядра); багатослівні n-грами не підтримуються форматом моделі, а леми зі списку lemma_overrides обробляються через буквальний пошук словоформи в корпусі. Відсіювання $174 = (94 + 92) - 3 \text{ lemma_overrides} - 9 \text{ лем, відсутніх у моделі.}$

Результати фільтрації. За результатами процедури з 174 однослівних seed-лем отримано 4 384 унікальних кандидатів, які класифіковано за чотирма статусами:

Таблиця И.6.1.

Розподіл кандидатів-маркерів за статусами автоматичної класифікації

Статус	К-сть	%	Критерій
auto_accept	3 044	69	висока оцінка, належність до однієї категорії
auto_accept_multivalent	134	3	висока оцінка, належність до двох категорій (score $\geq 0,55$ в обох)
manual_review	689	16	топ-20 на категорію для контролю якості, multivalent з нижчим score, кандидати у трьох і більше категоріях
rejected	517	12	composite score $< 0,40$

Правила розв'язання багатокатегорійності. Кандидати, що потрапили у три і більше категорій, не додаються до словника автоматично, а переходять у manual_review. Кандидати у двох категоріях зі score $\geq 0,55$ в обох додаються з контекстним правилом розв'язання, реалізованим через поле subject_level (див. Додаток И, розділ 4).

Результати ручної верифікації групи manual_review (689 кандидатів):

- 238 yes-кандидатів додано до фінального словника як нові standard-маркери;
- 97 yes-кандидатів збіглися з лемами, що вже потрапили до словника через seed або auto_accept_multivalent, і збережені без дублювання;
- 47 unclear-кандидатів додано з weak-вагою і обов'язковим контекстним підкріпленням;
- решта (307) відхилена.

Підсумок розширення. Фінальний розширений словник містить 3 518 маркерів (3 514 лем + 4 n-грами), що становить ≈ 13 -кратне зростання від стартового ядра в 262 маркери.

6.2. *Lemma_overrides*

Для лем, на яких бібліотека лематизації Stanza систематично виводить помилковий результат через омонімію або плюральні форми, передбачено механізм lemma_overrides: замість лематизованого пошуку виконується буквальний пошук словоформи. Приклад у таблиці И.6.2.

Таблиця И.6.2.

Приклад помилкових результатів лематизації Stanza

Лема	Категорія	Причина
рідня	CON	Stanza лематизує до «рідній»/«рідний»; буквальний пошук словоформ.
передова	UA-WAR	Stanza лематизує до прикметника «передовий», втрачаючи іменник «лінія фронту».
поминки	CON	Плюральна форма без однини; неконсистентний вихід лематизатора.

Перелік є відкритим: під час подальших ітерацій словника можуть додаватися інші лемми з обґрунтуванням і прикладом помилкового виходу Stanza.

6.3. *Класи маркерів і контекстні правила*

Усі маркери словника поділено на три класи надійності. Strong-маркери — високоспецифічні лемми, що спрацьовують самостійно без контекстного підтвердження (наприклад, «Голодомор», «вишиванка»); вага strong-лемми — 1,5; strong-n-грами — 2,0. Standard-маркери — типові лемми категорії, що спрацьовують самостійно з нижчою вагою; вага

— 1,0. Weak-маркери — низькоспецифічні лєми з численними нейтральними вживаннями; вимагають обов'язкового контекстного підтвердження у вікні ± 5 токенів; вага — 0,5.

Алгоритм присвоєння категорії абзацу: (1) перевірити strong-маркер, при знайденні — категорія = 1; (2) інакше — standard-маркер, при знайденні — категорія = 1; (3) інакше — weak-маркер з підтверджувальним контекстом, при знайденні обох — категорія = 1; (4) інакше — категорія = 0.

6.4. Службові поля

Окрім сімох бінарних рішень cat_X , для кожного абзацу обчислюється набір службових полів, які використовуються у діагностиці і якісному аналізі – таблиця И.6.3.

Таблиця И.6.3.

Набір службових полів, які використовуються у діагностиці і якісному аналізі

Поле	Зміст
intensity_X_weighted	Зважена сума ваг маркерів категорії X у межах абзацу. Кількісний показник щільності.
valence	Середня емоційна полярність слів у вікні ± 5 токенів від маркера (Tone-Dict-Ukrainian, шкала $[-2; +2]$). Описовий параметр; не підстава для присвоєння категорії.
carrier_auto	Тип голосу: dialogue_voice, first_person_narration, third_person_narration, unclear.
subject_level_auto	Рівень суб'єкта: individual, collective, institutional, mixed, unclear. Ключове поле для розмежування OPN/UA-FRE.
confidence	Інтегральний показник у $[0; 1]$; поєднує вагу маркерів, контекстну узгодженість, специфічність, частоту.

6.5. Попередня діагностика покриття seed-словника на корпусі

Перед запуском повного циклу автоматизованого кодування проведено попередню перевірку покриття seed-словника кодбуку на лематизованому корпусі.

Загальне покриття seed-лем у корпусі. Із 201 однослівної seed-леми кодбуку 199 (99,0%) мають принаймні одне входження у корпусі. За кількістю входжень леми розподілено на три діапазони:

- **loud** (понад 1 000 входжень) — 96 лем (48%);
- **mid** (100–1 000 входжень) — 67 лем (33%);
- **quiet** (менше 100 входжень) — 21 лема (11%).

Категорійне покриття активних лем (тих, що мають хоча б одне входження у корпусі) коливається від 84% (CON) до 93% (ENH), при середньому показнику 89% по всіх семи категоріях.

Леми, не виявлені у лематизованому корпусі. Дві леми категорії CON — «рідня» і «поминки» — очікувано пропускаються лематизатором Stanza через специфіку її лематизаційної моделі (омонімія, плюральні форми без однини). Обидві оброблюються через окремий механізм `lemma_overrides` — буквальний пошук словоформи замість лематизованого пошуку (див. розділ 6.2).

Підсумок. Попередня діагностика підтвердила готовність словника до етапу автоматичного розширення через `word2vec`-модель (див. розділ 6.1).

7. Довідково-інтерпретативний блок

Цей блок описує клас маркерів, що НЕ застосовуються у масовому автоматизованому кодуванні, а використовуються тільки у якісному аналізі ключових кейсів (30–50 найпоказовіших абзаців). Підстава для виокремлення — те, що відповідні структури реалізуються на латентному рівні, який потребує інтерпретативної реконструкції, а не лексичного збігу.

- **Наративні маркери:** сюжетний наслідок ціннісного вибору героя, ієрархізація персонажів за ціннісною шкалою, фокалізація як прихована точка зору автора, трансформація ціннісних орієнтацій персонажа протягом сюжету, моральні опозиції на рівні сюжетної структури.
- **Символічні маркери:** стійкі символи української культури (степ, могила, хата над Дніпром) у конотативно-ціннісному вживанні, образи-знаки колективної травми (вагон, табір, спалена хата), інтертекстуальні відсилання до Шевченка, Лесі Українки, Франка, Стуса як ціннісних авторитетів.

Спроба автоматизованого виявлення цих маркерів через лексичний словник принципово непродуктивна: один і той самий символ (наприклад, «степ») у різних контекстах активує протилежні ціннісні позиції (свобода козацтва >> UA-IDN, смерть переселенця >> UA-WAR, пустка байдужості >> без категорії). Тому для цього класу маркерів передбачено кваліфіковану якісну процедуру в межах КДА і наративного аналізу (підрозділ 2.2 основного тексту).

8. Каталог граничних випадків

8.1. Іронія, сарказм, постмодерністська дистанція

Ознаки: лапки на ціннісно навантаженому слові («наш справжній герой»), словесні маркери дистанції («нібито», «начебто», «так звана»), різкий стрибок між високою і низькою лексикою, контраст між заявленим і наслідком у сюжеті. У постмодерністській і експериментальній прозі іронічна дистанція до ціннісних маркерів є структурним прийомом.

Правило: категорія кодується як 1 (бо маркер у тексті присутній), у полі notes — позначка «іронія». Якісний аналіз окремо інтерпретує такі кейси як деконструкцію цінності.

8.2. Носій ціннісного судження (натор, персонаж, антагоніст)

Висловлювання може належати натору, головному герою, негативному персонажу, голосу у пам'яті або невластивому мовленню. Кодування фіксує лише наявність маркера у тексті, а не оцінку автора.

Правило: категорія кодується як 1 за наявністю маркера, незалежно від носія і авторської валентності. Авторська оцінка фіксується окремо у якісному аналізі; carrier_auto в автоматичному прогоні може зафіксувати неоднозначність як «unclear».

8.3. Цитати, інтертекст, метафоричні розширення

Цитата з іншого автора (Шевченка, Лесі Українки, народної пісні) або метафоричне розширення ключового слова можуть зберігати або втрачати ціннісне навантаження.

Правило: категорія кодується, якщо цитата інтегрована у наратив і ціннісне навантаження активне у контексті, або якщо метафоричне розширення зберігає прямий смисловий зв'язок з історичним явищем (наприклад, «це був голодомор для покоління»). Не кодується, якщо метафора явно віддалена від конкретного явища («холодна війна на роботі»). У спірних випадках — notes: «цитата» або «метафора».

8.4. Множинне кодування у складних абзацах

Один абзац може містити по 3–4 категорії одночасно. Приклад: «Мати готувала валізу для евакуації, дбайливо складаючи дитячі речі, і вклала туди вишиванку — щоб не забути, хто ми» одночасно отримує CON (родина), TRA (турбота), UA-WAR (евакуація), UA-IDN (вишиванка з ідентифікаційним контекстом). Множинне кодування — норма.

8.5. Сумніви 50/50 і незрозумілий абзац

У разі сумніву 50/50 щодо присутності категорії — ставити 1 з позначкою «сумнів» у notes. Це методично виправдано двома міркуваннями: мажоритарне голосування серед п'яти кодувальників статистично нейтралізує одиничні сумнівні голоси; втрата справжніх позитивів шкідливіша для агрегаційного аналізу, ніж обмежений шум хибних позитивів. Якщо абзац принципово незрозумілий навіть з контекстом (фрагмент діалогу без референта тощо) — проставляються нулі для всіх семи категорій або обирається «Нічого з зазначеного» з поміткою «незрозуміло».

Додаток К. Процедура ручного кодування абзаців

К.1. Процедура ручного кодування для валідації V1

Інструктивне навчання кодувальників. Інструктивна сесія для всіх п'яти кодувальників тривала приблизно 90 хвилин і охоплювала три блоки: (1) ознайомлення з категоріальною системою у функціональних термінах — без введення академічної термінології Шварца, з опорою на операційні визначення кодбука; (2) розбір пілотних прикладів кодування з аргументацією рішень — 8 показових абзаців, що ілюструють основні випадки активації кожної з 7 категорій; (3) тестове кодування 10–15 абзаців самими кодувальниками з подальшим обговоренням результатів і розв'язанням розбіжностей. Інструктивна сесія записувалася для забезпечення єдиного інструктажу всіх кодувальників і для можливості асинхронного навчання нових залучених кодувальників за необхідності.

Веб-інструмент розмітки. Технічно розмітка реалізована через спеціально розроблений веб-інструмент, який забезпечує однаковий формат подання тексту й контексту для всіх кодувальників, фіксує час прийняття рішення для кожного абзацу і дозволяє кодувальникам зберігати проміжний прогрес локально без передачі даних на сервер. Інструмент реалізовано як SPA (Single-Page Application) на React, з локальним сховищем у IndexedDB; вивантаження результатів виконується у форматі JSON для подальшого зведення у gold standards.

Технічна реалізація сліпого кодування. Кожен з п'яти кодувальників незалежно і у режимі сліпого кодування розмітив усі 100 абзаців корпусу V1 за сімома бінарними категоріями. Сліпе кодування технічно забезпечується через два окремі файли — coding-файл і metadata-файл, — які об'єднуються за ідентифікатором абзацу лише після завершення розмітки. Це усуває упередження кодувальника, пов'язане з очікуваннями щодо динаміки (наприклад, очікування зростання UA-WAR у 2022–2025 рр.) і робить ручне кодування методологічно більш строгим.

К.2. Промпт для звернення до API LLM. Структура промпта

""Ви — досвідчений експерт-кодувальник у соціології ціннісних орієнтацій (теорія Шварца) та аналізі української художньої прози. Ваше завдання — провести аксіологічну розмітку коротких абзаців з художніх творів за 7 категоріями цінностей.

ВАЖЛИВО:

- Кодуєте ТІЛЬКИ те, що видно у тексті — не додумуйте підтекст і не аналізуйте «що мав на увазі автор».
- Категорії незалежні: один абзац може мати 0, 1 або кілька активних категорій.
- Якщо текст нейтральний (опис побуту, природи без оцінних відтінків) — це нормально, ставите порожній список.
- Якщо сумнівно (50/50) — краще поставити категорію і позначити сумнів у rationale.

СИМ КАТЕГОРІЙ:

◆ OPN — Відкритість до змін, індивідуальна свобода

Про окрему людину: свобода вибору, новизна, ризик, творчість, мрія, насолода, пригода, кохання (романтичне), втеча від обмежень.

ВАЖЛИВО: тільки про індивіда. «Свобода народу» — це не OPN, а UA-FRE.

Маркери: свобода, вибір, мрія, ризик, пригода, нове, кохання, втекти, звільнитися, творити, експериментувати, насолода.

✓ ТАК: «Він купив квиток і вирушив, не озираючись» — індивідуальна свобода вибору.

✓ ТАК: «Вона мала право на своє щастя і свої помилки» — індивідуальна свобода.

✗ НІ: «Народ виборів незалежність» — це колективне >> UA-FRE.

◆ CON — Збереження, родина, традиції

Про родину, дім, традиції, обряди, повагу до старших, безпеку, порядок. Стабільність, конформність, оглядка на громаду.

Маркери: родина, сім'я, мати, батько, дім, хата, рідня, обов'язок, традиція, звичай, обряд, віра, церква, рушник, весілля, поминки, спадщина, повага, шаноба.

✓ ТАК: «Мати поправила рушник на іконі, як завжди в цій хаті» — родина, традиція, дім.

✓ ТАК: «Що скажуть люди? Так не годиться» — оглядка на громаду.

✗ НІ: «Вишиванка стала для нього символом українства» — це UA-IDN.

✗ НІ: «Він зайшов додому, повісив куртку» — побутове «дім», не ціннісне.

◆ TRA — Турбота про інших, доброта, справедливість

Допомога іншим, співчуття, милосердя, дружба, самопожертва, справедливість. Любов як турбота (НЕ романтична).

Маркери: допомога, врятувати, віддати, ділитися, дружба, друг, співчуття, милосердя, доброта, прощення, відданість, справедливість, чесність, людяність, природа (як об'єкт захисту)

✓ ТАК: «Бабуся віддала їй останній шматок хліба, сама лягла голодна» — самопожертва.

✓ ТАК: «Він підняв пораненого з вулиці і доніс до лікарні» — допомога.

Х НІ: «Вони палко цілувалися всю ніч» — це OPN (романтика).

◆ ENH — Успіх, влада, престиж

Про індивідуальний успіх, перемогу над суперником, статус, владу, гроші, визнання. Індивідуальна вершина.

ВАЖЛИВО: тільки про індивіда. «Перемога народу» — UA-FRE/UA-WAR.

Маркери: успіх, перемога, виграти, тріумф, амбіція, кар'єра, гроші, багатство, влада, керувати, наказати, престиж, статус, авторитет, найкращий, контроль, домінування, конкуренція.

✓ ТАК: «Він добився посади директора, купив дороге авто» — статус, гроші.

✓ ТАК: «Перемогти в цьому турнірі — означало стати першим» — індивідуальна перемога.

Х НІ: «Україна виграла війну за свободу» — колективне >> UA-FRE/UA-WAR.

◆ UA-IDN — Українська ідентичність, культурна самотність

Усвідомлення українства, мова, культура, історія, символіка. Самоідентифікація з нацією, повернення до коренів.

Маркери: українець, україномовний, український, козак, гетьман, Січ, Україна, нація, народ, мова рідна, материнська мова, Київ/Львів/Харків (як символи), Шевченко, Леся Українка, Франко, вишиванка, рушник (з культурним контекстом), писанка, тризуб, батьківщина, ідентичність.

✓ ТАК: «Я — українець, і це найголовніше» — пряма самоідентифікація.

✓ ТАК: «Він повернувся до української мови, бо це мова його прадідів» — мова як ідентифікація.

Х НІ: «Він жив у Києві на проспекті Перемоги» — Київ як топонім, не символ.

◆ UA-WAR — Війна, репресії, історичні катастрофи

Досвід війни (ВВ, АТО, повномасштабне вторгнення 2022), репресії, Голодомор, Чорнобиль, переселення, втрата дому. Особистий або колективний досвід катастрофи.

Маркери: війна, фронт, окоп, бомба, обстріл, ракета, тривога, окупація, окупант, ворог, біженець, переселенець, евакуація, загиблий, втрата, Голодомор, голод, репресія, ГУЛАГ, Сибір, депортація, Чорнобиль, Бахмут, Маріуполь, Майдан (як трагедія), травма, пам'ять.

✓ ТАК: «Вони вибралися з оточеного міста через гуманітарний коридор» — переміщення під час війни.

✓ ТАК: «Бабуся переписала рецепт хліба, який вона їла у тридцять третьому» — Голодомор.

✗ НІ: «Холодна війна на роботі» — метафора, не справжня війна.

◆ UA-FRE — Свобода нації, спротив, революція

Колективна свобода: національна незалежність, спротив окупації, революційні події (Майдан 1990, Помаранчева 2004, Революція Гідності 2014), боротьба за державність. Колективний суб'єкт.

Маркери: визволення, гідність, Майдан, революція гідності, спротив, опір, протистояння, повстання, суверенітет, незалежність, барикада, віче, європейський вибір, «не дамося», «не здамося», «ми вистояли», голос народу.

✓ ТАК: «Ми стояли на Майдані три місяці і вистояли» — колективна революційна дія.

✓ ТАК: «Український народ не здається перед загарбником» — колективний суб'єкт.

✗ НІ: «Він хотів свободи — щоб ніхто не вказував йому» — індивідуальна >> OPN.

ПРАВИЛА РОЗМЕЖУВАННЯ СУМІЖНИХ КАТЕГОРІЙ:

1. «Свобода» особистості >> OPN; «свобода народу» >> UA-FRE.
2. «Перемога» особиста >> ENH; «перемога народу» >> UA-FRE/UA-WAR.
3. «Традиція» побутова/родинна >> CON; українська національна (вишиванка з контекстом) >> UA-IDN. Часто обидві.
4. Якщо неясно (50/50) — ставити обидві категорії.

ОСОБЛИВІ СИТУАЦІЇ:

- Іронія/сарказм (лапки навколо цінного слова, «нібито», «начебто»): категорія все одно активна, у rationale зазначити «іронія».
- Згадка минулого або мрії: категорія активна, навіть якщо ситуація сюжетно віддалена.
- Слова негативного персонажа: кодуємо так само — «чи присутнє у тексті», не «чи схвалює автор».
- Якщо абзац нейтральний (опис побуту без оцінних маркерів) — порожній список категорій. ""

Спрощений промпт для Gemini Pro — без category_check,

щоб зменшити обсяг output (thinking mode у Gemini витрачає багато токенів).

Категорії все одно отримуються через "analysis" + "categories".

USER_PROMPT_GEMINI_TEMPLATE = ""Аналізуйте абзаци і дайте відповідь у JSON.

КОНТЕКСТ ДО:

{context_before}

ОСНОВНИЙ ТЕКСТ:

{text}

КОНТЕКСТ ПІСЛЯ:

{context_after}

ЗАВДАННЯ:

1. Коротко проаналізуйте ОСНОВНИЙ текст (1-3 речення): про що, які ціннісні маркери.
2. Виберіть АКТИВНІ категорії з 7: OPN, CON, TRA, ENH, UA-IDN, UA-WAR, UA-FRE.
3. Дайте JSON.

ФОРМАТ ВІДПОВІДІ — суворо JSON, нічого більше:

```
```json
```

```
{
 "analysis": "коротка інтерпретація",
 "categories": ["активні", "категорії"],
 "confidence": "high|medium|low",
 "notes": "якщо є особливості — зазначити, або порожнє"
}
```

```
```
```

Якщо жодна категорія не активна — "categories": [].""

Додаток Л. Попарна α -матриця узгодження кодувальників

Таблиця Л.1.

Попарна α -матриця узгодження між 8 кодувальниками (5 людей + 3 LLM)

| Кодувальник | A | L | N | T | V | Claude Opus 4.7 | Gemini 2.5 Pro | GPT-4o |
|-----------------|-------|-------|-------|-------|-------|-----------------|----------------|--------|
| A | 1,000 | 0,384 | 0,291 | 0,146 | 0,362 | 0,439 | 0,384 | 0,452 |
| L | 0,384 | 1,000 | 0,303 | 0,183 | 0,438 | 0,307 | 0,256 | 0,362 |
| N | 0,291 | 0,303 | 1,000 | 0,119 | 0,379 | 0,390 | 0,201 | 0,333 |
| T | 0,146 | 0,183 | 0,119 | 1,000 | 0,130 | 0,202 | 0,170 | 0,309 |
| V | 0,362 | 0,438 | 0,379 | 0,130 | 1,000 | 0,362 | 0,272 | 0,458 |
| Claude Opus 4.7 | 0,439 | 0,307 | 0,390 | 0,202 | 0,362 | 1,000 | 0,561 | 0,634 |
| Gemini 2.5 Pro | 0,384 | 0,256 | 0,201 | 0,170 | 0,272 | 0,561 | 1,000 | 0,574 |
| GPT-4o | 0,452 | 0,362 | 0,333 | 0,309 | 0,458 | 0,634 | 0,574 | 1,000 |

Джерело: власні розрахунки за VI-вибіркою; групи кодувальників: 5 людей та 3 LLM-моделі.

Додаток М. Метрики precision / recall / F1

Таблиця М.1.

Метрики precision / recall / F1 за категоріями і трьома gold standards (n=94 абзаци)

| Категорія | Gold | TP | FP | FN | TN | Precision | Recall | F1 |
|---------------|----------|----|----|----|----|-----------|--------|--------------|
| OPN | Human | 5 | 32 | 6 | 51 | 0,135 | 0,455 | 0,208 |
| | LLM | 12 | 25 | 16 | 41 | 0,324 | 0,429 | 0,369 |
| | Expanded | 7 | 30 | 13 | 44 | 0,189 | 0,35 | 0,246 |
| CON | Human | 0 | 24 | 3 | 67 | 0 | 0 | — |
| | LLM | 3 | 21 | 7 | 63 | 0,125 | 0,3 | 0,176 |
| | Expanded | 1 | 23 | 7 | 63 | 0,042 | 0,125 | 0,062 |
| TRA | Human | 1 | 21 | 4 | 68 | 0,045 | 0,2 | 0,074 |
| | LLM | 1 | 21 | 11 | 61 | 0,045 | 0,083 | 0,059 |
| | Expanded | 1 | 21 | 9 | 63 | 0,045 | 0,1 | 0,063 |
| ENH | Human | 1 | 26 | 7 | 60 | 0,037 | 0,125 | 0,057 |
| | LLM | 5 | 22 | 17 | 50 | 0,185 | 0,227 | 0,204 |
| | Expanded | 3 | 24 | 16 | 51 | 0,111 | 0,158 | 0,13 |
| UA-IDN | Human | 0 | 13 | 4 | 77 | 0 | 0 | — |
| | LLM | 0 | 13 | 6 | 75 | 0 | 0 | — |
| | Expanded | 0 | 13 | 7 | 74 | 0 | 0 | — |
| UA-WAR | Human | 0 | 18 | 2 | 74 | 0 | 0 | — |
| | LLM | 3 | 15 | 9 | 67 | 0,167 | 0,25 | 0,2 |
| | Expanded | 1 | 17 | 2 | 74 | 0,056 | 0,333 | 0,095 |
| UA-FRE | Human | 0 | 10 | 1 | 83 | 0 | 0 | — |
| | LLM | 1 | 9 | 5 | 79 | 0,1 | 0,167 | 0,125 |
| | Expanded | 0 | 10 | 1 | 83 | 0 | 0 | — |

Примітка. Типе (—) у колонці F1 означає, що TP=0 і F1 неможливо обчислити (precision і recall дорівнюють нулю). Розрахунок precision / recall / F1 виконано на 94 абзацах V1 після вилучення 6 випадків із колізіями ідентифікаторів під час зіставлення автоматичного й ручного кодування; ці вилучення не змінюють стратифікаційної логіки V1.

Додаток Н. Попарне співпадіння трьох gold standards між собою

Таблиця Н.1.

Попарне співпадіння трьох gold standards між собою (overall agreement, %)

| Пара gold | | OPN | CON | TRA | ENH | UA-IDN | UA-WAR | UA-FRE | Overall |
|-----------------------|---|-------|-------|-------|-------|--------|--------|--------|--------------|
| Human LLM | - | 79,0% | 92,0% | 92,0% | 81,0% | 93,0% | 88,0% | 95,0% | 88,6% |
| Human Expanded | - | 88,0% | 95,0% | 95,0% | 87,0% | 96,0% | 98,0% | 100,0% | 94,1% |
| LLM Expanded | - | 91,0% | 97,0% | 97,0% | 94,0% | 97,0% | 90,0% | 95,0% | 94,4% |

Додаток П. Розширений якісний аналіз кейсів

Кейс П.1. Білик Іван. «Не дратуйте грифонів», 1993 (блок Е, 1991–2003). Діалогова репліка з 28 слів (para_id=s008_p0046; автокод: UA-IDN + ENH):

«— І як же ти з твоїм Фукідідом угадаєш, хто дужчий? Досі було легко. Скіфи могутніші за тих агафірсів, і ваша хитрість могла принести вам користь. А тепер?».

На лексико-семантичному рівні UA-IDN активується через етнічними **скіфи** й **агафірси** та апеляцію до **Фукідида** — лексику канонічного шару історико-археологічної рамки ранньонезалежницького дискурсу про дохристиянські, дотатарські, доросійські корені українських теренів; ENH активується через семантику сили й переваги (*могутніші, дужчий, користь*). На наративному рівні маркер UA-IDN тут — мовлення персонажа в межах історичного діалогу, а не пряма авторська декларація. На дискурсивному рівні фрагмент є типовим для блоку Е у першому зрізі, де UA-IDN сягає піка 17,56% : історична проза 1990-х систематично активує ідентичнісні маркери через апеляцію до античних і середньовічних коренів, обираючи стратегію легітимації українського «себе» через дистанційну, дотатарську темпоральність.

Кейс П.2 Новік Андрій. «Медальйон трьох змій», 2022 (блок В, 2022–2025). Наративний абзац із 86 слів (автокод: OPN з інтенсивністю 4,0, ENH з інтенсивністю 2,0):

«Хлопець вдячно кивнув та попростував за рештою команди. Крок за кроком вони протиснулися крізь вузький коридор до драбини, що виводила на спину “Стріли”. Першим виліз Герберт, подав руку Анні, далі за нею... Едмунд, не звиклий до довгих подорожей, під водою й поготів...».

На лексико-семантичному рівні OPN активується через лексику руху, простору й мандрів (*повітря, подорож, картинка*), а ENH — через ономастичну насиченість іменами персонажів-агентів (Герберт, Анна, Влад, Фердинанд, Едмунд, Мечислав), які в межах словникової евристики кодуються як маркери досягнення. На наративному рівні позиція об'єктивного третьоособового наратора фіксує колективну дію команди в жанровій конвенції пригоди / фентезі без прямої ціннісної рефлексії. На дискурсивному рівні фрагмент **частково підтверджує H2s у блоці В** (зростання ENH з 15,64% до 19,31% між першим і четвертим зрізами), але водночас ілюструє важливе обмеження інтерпретації: значна частка автоматичної активації ENH у жанровій белетристиці походить не з власне моделей особистого успіху чи

кар'єри, а з ономастично-агентної лексики пригодницько-фентезійної конвенції. Тобто частина зростання ENH у блоці В — це не модернізаційний зсув цінностей самоствердження в шварцівському сенсі, а зростання частки самого жанрового підтипу, де ця лексика структурно щільніша. Це додатковий аргумент на користь обережного формулювання висновку 4.3 — H2s «частково підтримується», а не «підтверджується».

Кейс П.3. Рафєєнко Володимир. «Петрикор — запах землі після дощу», 2023 (блок Е, 2022–2025). Наративний фрагмент із 91 слова (автокод: TRA + ENH, **але знов не UA-IDN і не UA-FRE**):

«...Її чоловік, її коханий Сашко, Ангел, Крішна-на-Вишні, її Шовковичний Господар не міг приїхати сюди до неї після входу в місто російських бойовиків. Крім всіляких інших причин, це було просто небезпечно хоч би вже й тому, що він реально російською спілкуватися був абсолютно не здатний. Кілька разів з поваги до її батьків намагався за спільним столом, але виходило так погано, що навіть Марійчин батько попросив Сашка продовжувати спілкуватися рідною мовою».

Це другий кейс F1=0-типу для UA-IDN. Фрагмент організовано довкола неможливості російськомовного спілкування як межі особистої й сімейної інтимності в умовах війни, але словник UA-IDN не активується; натомість TRA активується через лексику любові й поваги (коханий, з поваги), ENH — імовірно через сакралізаційні епітети (Ангел, Шовковичний Господар). На наративному рівні третьоособова фокалізація крізь жіночу свідомість фіксує мовну приналежність як ціннісну межу між світами. На дискурсивному рівні фрагмент особливо інформативний з огляду на біографію автора: Рафєєнко як білінгвальний письменник, що сам перейшов на українську після 2022 р., конструює персонажа з симетричним досвідом мовної невступу в чужий світ — і саме ця конструкція є ціннісною позицією, яку автоматичне кодування не реєструє. Кейс підтримує тезу 4.5 про зсув від лексично-маркованого до наративно-маркованого режиму вираження ідентичнісних позицій у пізніх зрізах.

Кейс П.4. Лис Володимир. «Вифлєєм», 2022 (блок А, 2022–2025). Діалоговий фрагмент із 62 слів (автокод: жодної з семи категорій не активовано):

«— Я хочу сказати, що ти заслуговуєш на значно більше, ніж маєш, — Клаудія посміхнулася зовсім по-іншому, але найперше — заохочувально. — І я хочу приймати в іншому палаці набагато поважніших матрон, ніж досі. Ти все зможеш, бо ти розумний. Не лише

сильний, як і належить справжньому воїну, а й розумний. Ти будеш там, де тобі й належить бути, Публій Сульпіцій Квіріній!».

Цей кейс ілюструє поєднання **Межі 1** і **Межі 4**, описаних у підрозділі 2.2: автомат не активує жодної категорії, попри те що репліка щільно насичена ціннісною семантикою. На лексико-семантичному рівні присутні маркери, потенційно близькі до ENH (*заслугуєш, поважніших, справжньому воїну, належить бути*) і до CON (*належить бути* — конформність ролі), однак вони сформульовані в архаїчно-урочистій модальності давньоримського наративу, яка не перетинається зі словниковими лемами, налаштованими на сучасну українську ціннісно-маркерну лексику. На наративному рівні дружня репліка Клаудії, звернена до Публія Сульпіція Квірінія (історичної постаті римського намісника, відомого з євангельського наративу про перепис), функціонує як «жіноча мобілізаційна» риторика статусу й амбіцій. На дискурсивному рівні цей фрагмент демонструє важливе обмеження автоматизованого підходу: ціннісно-насичений матеріал, що належить до стилістично-маркованого піджанру історико-релігійної прози, може цілком не реєструватися словником, налаштованим на сучасну ціннісно-маркерну лексику. Це додатково підтримує висновок підрозділу 4.6 про доцільність подальшого використання контекстуалізованих векторних моделей для випадків нетипової стилістичної модальності.

Додаток Р. Попарні порівняння часток активних абзаців між хронологічними зрізами

Таблиця Р.1.

**Попарні порівняння часток активних абзаців між хронологічними зрізами:
p-value з поправкою FDR / Cohen's h**

| Категорія | 91–03 ↔ 04–13 | 04–13 ↔ 14–21 | 14–21 ↔ 22–25 | 91–03 ↔ 22–25 |
|-----------|-----------------|-----------------|-----------------|-----------------|
| OPN | <0,001 / –0,038 | <0,001 / +0,019 | <0,001 / –0,022 | <0,001 / –0,041 |
| CON | <0,001 / –0,090 | <0,001 / +0,021 | <0,001 / –0,058 | <0,001 / –0,127 |
| TRA | <0,001 / –0,040 | <0,001 / +0,009 | 0,099 / –0,004 | <0,001 / –0,036 |
| ENH | <0,001 / –0,035 | <0,001 / +0,024 | <0,001 / –0,024 | <0,001 / –0,035 |
| UA-IDN | <0,001 / –0,134 | 0,003 / –0,006 | <0,001 / –0,062 | <0,001 / –0,202 |
| UA-WAR | <0,001 / –0,093 | <0,001 / +0,051 | <0,001 / –0,032 | <0,001 / –0,074 |
| UA-FRE | <0,001 / –0,071 | 0,129 / +0,003 | <0,001 / –0,019 | <0,001 / –0,087 |

Примітка. Напрямок h відповідає зміні від першого періоду пари до другого; від'ємне значення означає спад активності, додатне — зростання. Значення розраховані за агрегованими пропорціями, виведеними з ноутбука 03; p-value подано після FDR-корекції. Джерело 03_trends_pairwise_comparisons.csv

Додаток С. Рисунки декомпозиції та Z-статистик

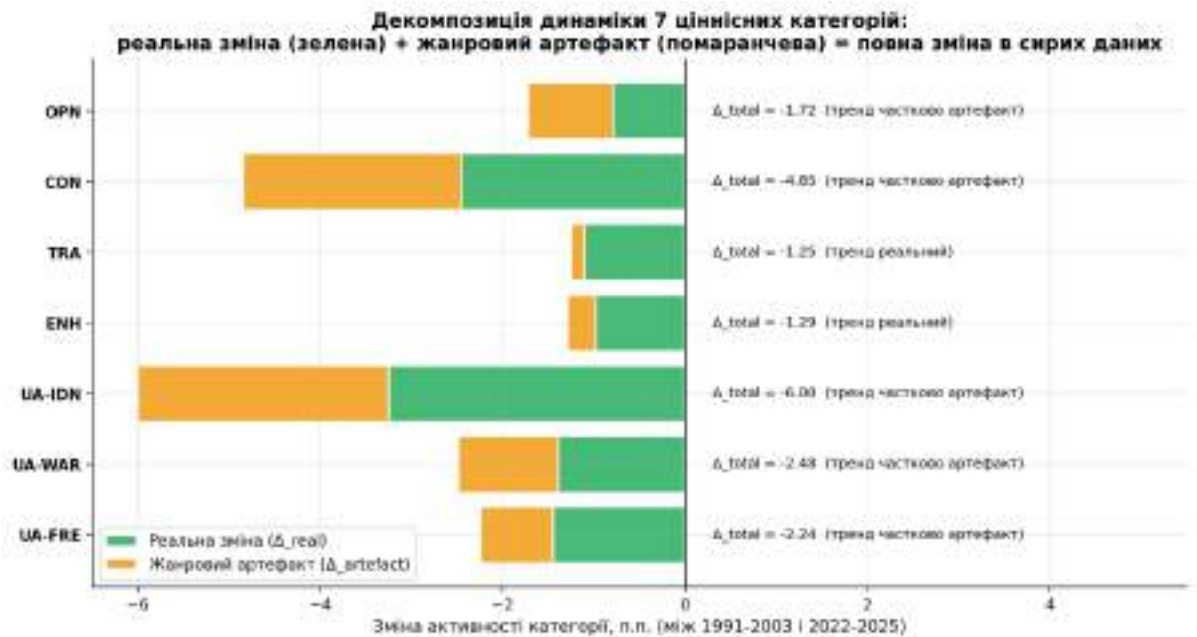


Рис. С.1. Декомпозиція динаміки семи ціннісних категорій: власне ціннісна зміна та жанровий артефакт

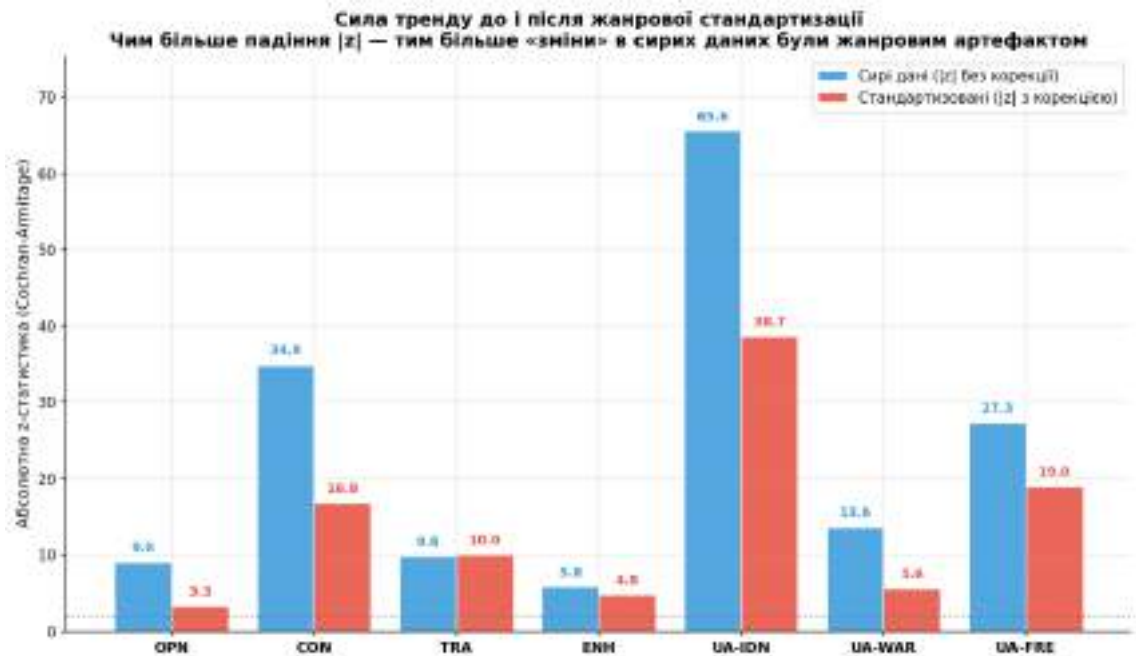


Рис. С.2. Абсолютні Z-статистики Cochran–Armitage до і після жанрової стандартизації

Додаток Т. Триангуляція корпусних результатів з даними World Values Survey і European Values Study. Опорні точки національних опитувань.

Цей додаток містить технічні деталі підрозділу 4.5: повну карту операціональних відповідностей між категоріями кодбука й шкалами WVS/EVS, методологічні пояснення двох розв'язок (Y003_simple замість Y003 та EVS-secondary fallback), таблицю композитного WVS-показника, таблицю з репрезентативними даними національних опитувань та два додаткові рисунки.

Т.1. Карта операціональних відповідностей категорій кодбука та WVS/EVS змінних

Таблиця Т.1

Повна карта операціональних відповідностей: WVS main + secondary, EVS-відповідники, напрямок нормалізації

| Кат. | Концепт | WVS main +
secondary | EVS відповідник | Напрямок
нормалізації |
|------------|--|---|--|-------------------------------------|
| OPN | Відкритість до змін; автономія індивіда | main: Y003_simple = mean(Y021A, Y021B, Y021C); sec: Y003 (готовий Welzel), Y023 | Y003_simple обчислюється з child qualities A029 (Independence), A034 (Imagination), A042 (Obedience, інв.) | higher =
greater
autonomy |
| CON | Збереження / традиційно-консервативні орієнтації | main: Y010 SACSECVAl (інв.); sec: F063 (важливість Бога), E018 (повага до авторитету, інв.) | F063 — fallback (Welzel Y010 у EVS відсутній); F063 у EVS-Trend File: V123/V129/v63 | higher =
greater
conservatism |
| TRA | Само-трансцендентність; міжособистісна довіра | main: A165 «Most people can be trusted»; sec: Y022, B008 | A165 — те ж питання у EVS (V66/V62/v31); B008 відсутнє у EVS 2008, доступне у EVS 1999/2017 | higher =
greater
trust |

| | | | | | |
|--------------------|---|---|---|--|---|
| ENH | Самоствердження;
матеріалістичні
орієнтації | main: Y001 Post-
Materialist 12-item
(інв.); sec: E035
(рівність доходів) | E035 — fallback
(Welzel Y001 у EVS
відсутній); E035 у
EVS: O18/V198/v106 | higher
greater
materialism | = |
| UA-
IDN | Українська
національна
ідентичність | main: G006 «How
proud of nationality»
(інв.); sec: Y011B | G006 — те ж питання
у EVS
(V255/V256/v170) | higher
greater pride | = |
| UA-
WAR | Травма війни;
військова тема | main: E069_02
Confidence Armed
Forces (інв.); sec:
Y014A (Welzel, інв.) | E069_02 — те ж
питання у EVS
(V201/V206/v116) | higher
greater
confidence | = |
| UA-
FRE | Свобода;
колективна
суб'єктність | main: Y024 Welzel
Voice subindex; sec:
E117 (підтримка
демократії, інв.),
E023 (інтерес до
політики, інв.) | E117 — fallback (Y024
у EVS відсутня); E117
у EVS:
V219/V228/v148 | higher
greater
pro-
democracy
stance | = |

Т.2. Дві методологічні розвилки

Розвилка 1. Y003_simple замість готового WVS Y003. Готовий індекс Y003 у WVS Trend File реалізує неопубліковане Welzel-перетворення (за «Freedom Rising», 2013), яке не відтворюється в EVS Trend File без оригінального GESIS-syntax коду. Діагностика показала, що готовий Y003 і просте середнє $\text{mean}(Y021A, Y021B, Y021C)$ розходяться по чотирьох українських хвилях на величини 0,19–0,41 (raw 0–1 шкала). При цьому Y021A повністю збігається з $\text{share}(A029=1)$, тобто sub-індекси є простими частками. Для прозорості й коректного перенесення в EVS, як основну змінну для OPN використано $Y003_simple = (Y021A + Y021B + Y021C) / 3$. Готовий Y003 збережено як secondary.

Розвилка 2. EVS-secondary fallback. Для трьох категорій (CON, ENH, UA-FRE) main-змінні є Welzel-індексами (Y010, Y001, Y024), яких немає у публічному EVS Trend File. Для цих категорій EVS-точки розраховано через перші доступні допоміжні (secondary) змінні: F063 для CON, E035 для ENH, E117 для UA-FRE. У зведенні «direction_match_extended» WVS і EVS аналізуються роздільно, а не як одна серія — це методологічно чистіше, ніж змішування різних метрик.

Т.3. Композитний WVS-показник (main + усі secondary)

Для перевірки чутливості висновків до вибору main-змінної розраховано додатково композитний показник — середнє нормалізованих значень main і всіх secondary всередині WVS (для тих категорій, де secondary доступні). Композитний direction_match (Таблиця Т.2) збігається з main-only для шести категорій з семи; розбіжність зафіксовано лише для CON, де main-only дає DIVERGENT, а композит — CONVERGENT. Це відображає реальну внутрішню неузгодженість концепту CON (Y010 стабільна, F063 росте, E018 падає).

Таблиця Т.2

Композитний WVS-показник vs main-only по семи категоріях

| Категорія | Корпус Δ , п.п. | WVS comp.
перший | WVS comp.
останній | Composite
direction |
|-----------|------------------------|---------------------|-----------------------|------------------------|
| OPN | −0,8 | 46,3 | 50,9 | DIVERGENT |
| CON | −2,4 | 61,7 | 59,6 | CONVERGENT |
| TRA | −1,1 | 45,8 | 45,8 | FLAT |
| ENH | −1,0 | 67 | 61,6 | CONVERGENT |
| UA-IDN | −3,2 | 60,1 | 72,8 | DIVERGENT |
| UA-WAR | −1,4 | 60 | 63,5 | DIVERGENT |
| UA-FRE | −1,5 | 44,8 | 46,4 | DIVERGENT |

Таблиця Т.3.

Хвилі WVS і EVS для України: обсяги вибірок

| Архів | Wave | Рік | n (приблизно) | Метод | Примітки |
|-------|------|------|---------------|-------------|------------------------|
| WVS | 3 | 1996 | ≈ 2 800 | F2F | Найбільша вибірка |
| WVS | 5 | 2006 | ≈ 1 000 | F2F | — |
| WVS | 6 | 2011 | ≈ 1 500 | F2F | Найменша точність |
| WVS | 7 | 2020 | ≈ 1 300 | F2F + Mixed | Частково під час COVID |
| EVS | 3 | 1999 | ≈ 1 200 | F2F | — |
| EVS | 4 | 2008 | ≈ 1 500 | F2F | B008 відсутня |
| EVS | 5 | 2017 | ≈ 1 400 | F2F | — |

Т.4. Додаткові рисунки

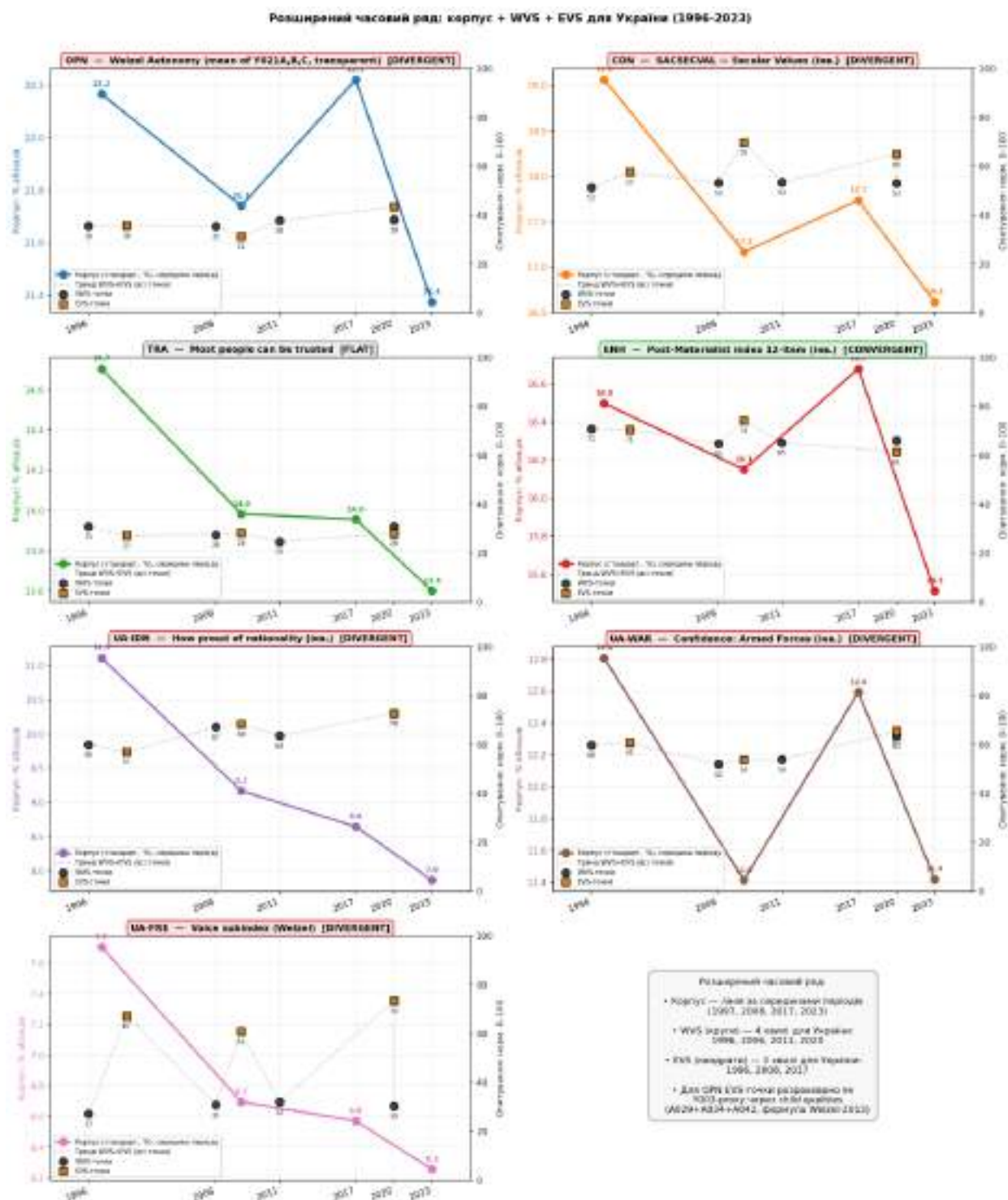


Рисунок Т.1 Розширений часовий ряд: корпус (середини періодів) + WVS- і EVS-точки для України, 1996–2023 рр.

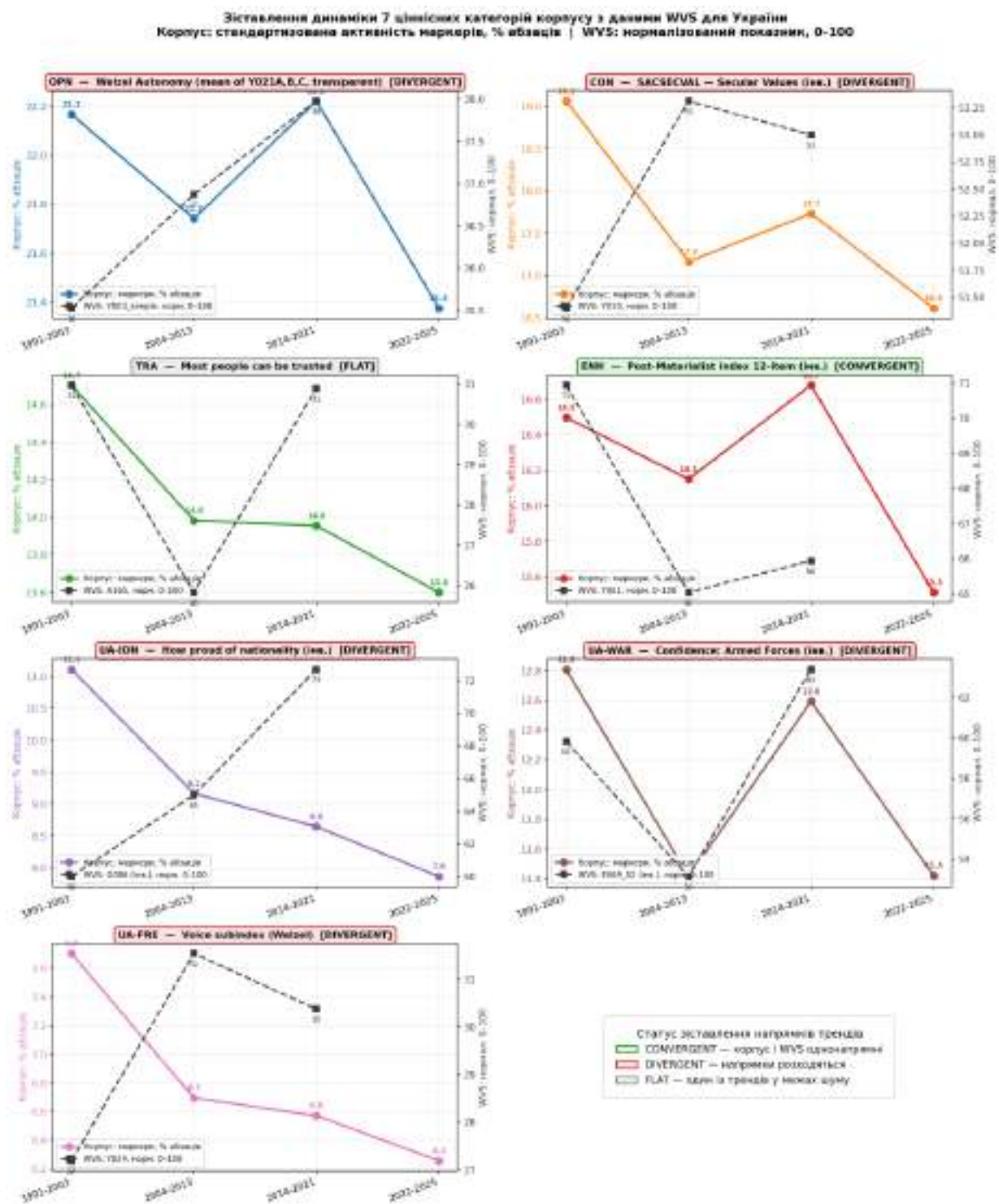


Рисунок Т.2 Зіставлення динаміки 7 ціннісних категорій корпусу з даними WVS для України (по 4 хронологічних зрізах)

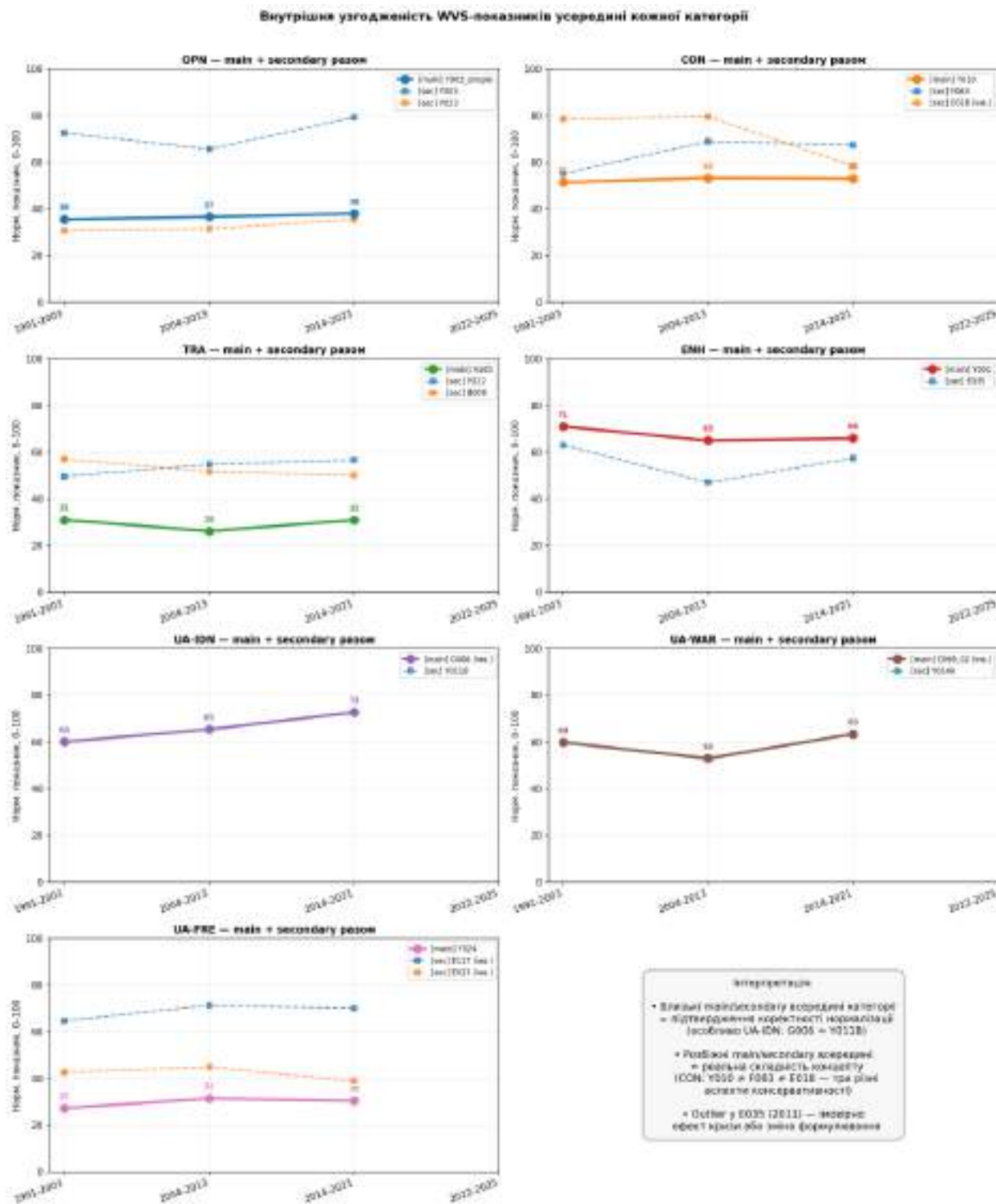


Рисунок Т.3 Внутрішня узгодженість main- і secondary-показників WVS усередині кожної категорії

Т.5. Опорні точки національних опитувань 2022–2025 рр.

Через відсутність WVS Wave 8 і EVS-вимірювань для четвертого хронологічного зрізу контекст 2022–2025 рр. (підрозділ 4.5) реконструюється за репрезентативними національними опитуваннями. Нижче — опорна таблиця за чотирма категоріями (TRA і ENH не мають співмірних національних індикаторів у публічному доступі).

Таблиця Т.4

Опорні точки національних опитувань 2022–2025 рр.

| Кат. | Індикатор | Значення | Джерело, дата |
|--------|--|-------------------------------|---|
| UA-IDN | Самоідентифікація «громадянин України» | 80% (2014: 51%, 2022: 60%) | КМІС, жовтень 2024, n=2004 |
| UA-IDN | Етн. українці з однією національністю | 84% (2017: 77%) | Razumkov, червень 2024, n=2008 |
| UA-WAR | Довіра до ЗСУ | 92% | КМІС грудень 2024; Razumkov листопад 2025 |
| CON | Частка віруючих (динаміка 2021–2024) | 68 >> 74 >> 70,5 >> 68% | Razumkov, жовтень 2024, n=2017 |
| CON | Довіра до Церкви | 62,5% (макс. 72,5% у 2010 р.) | Razumkov, жовтень 2024 |
| UA-FRE | Україна за повною згодою з демократією | 6-те місце з 41 країни | Gallup Int. / Рейтинг, EOY 2023 |
| UA-FRE | Підтримка вступу до ЄС | 75% | IRI, липень 2025, n=2400 |