Дослідницький центр «Прізма»
Prism Research Center

Методологічна записка

Тріаж для перевантаженої черги

Практичний метод тріажу для огляду великих обсягів публічних повідомлень

Вересень 2026

Методологічна записка

Тріаж для перевантаженої черги

Практичний метод тріажу для огляду великих обсягів публічних повідомлень

Автор	Дослідницький центр «Прізма», дослідницька команда (служба моніторингу)
Опубліковано	Вересень 2026
Охоплений період	3 березня 2025 – 31 травня 2026 (65 тижнів)
Ліцензія	Вільно читати, завантажувати й використовувати повторно. © Дослідницький центр «Прізма».

Рекомендоване цитування

> Дослідницький центр «Прізма» (2026). *Тріаж для перевантаженої черги: практичний метод тріажу для огляду великих обсягів публічних повідомлень*. Методологічна записка. Дослідницький центр «Прізма». https://prismresearch.org.ua/uk/publications/

Дослідницький центр «Прізма»
Prism Research Center

Контакти
research@prismresearch.org.ua · prismresearch.org.ua

Анотація

Упродовж 65-тижневого періоду високого навантаження (3 березня 2025 – 31 травня 2026) служба моніторингу «Прізми» отримала 18 240 публічних повідомлень приблизно з 30 відібраних джерел, тоді як можливості першого огляду становили близько 280 еквівалентів повідомлень на тиждень. У найнапруженіші тижні надходження перевищували можливості першого огляду, і повідомлення, що надійшли на початку пікового тижня, могли днями чекати на перший огляд. Ми розробили метод тріажу з трьох частин: (1) ранжування вхідних повідомлень за оціненою ймовірністю того, що їхній огляд приведе до дій, за допомогою каліброваної моделі на основі публічних метаданих і текстових ознак; (2) прогнозування обсягу високопріоритетних повідомлень на наступний тиждень для планування годин роботи волонтерів; (3) розподіл тижневого ресурсу на рецензування між комірками «регіон – категорія» зі збереженням мінімального покриття критичних комірок. Під час семитижневого проспективного тіньового тестування (13 квітня – 31 травня 2026), у якому позначки фіксували до будь-яких операційних дій, повідомлення з верхніх 15% були придатними до дій у 28,6% випадків, що в 3,2× перевищує базовий рівень 8,9% (повнота 48,2%). Моделювання черги на всьому архіві показало, що огляд у ранжованому порядку скоротив медіанний час до першого огляду придатних до дій повідомлень з 56 годин до 1 години. Усі вхідні дані були публічними, а всі рішення ухвалювали аналітики; модель лише ранжує та надає рекомендації.

Зміст

1 Проблема
2 Дані та позначки
3 Модель
4 Оцінювання моделі
5 Проспективне тіньове тестування
6 Прогнозування навантаження

7 Розподіл ресурсу на рецензування
8 Операційний ефект (моделювання)
9 Інструменти
10 Ролі
11 Обмеження

1 Проблема

За охоплений період надходження становили в середньому 281 повідомлення на тиждень (медіана 272, пік 942). Це сирі кількості отриманих повідомлень до вилучення майже дублікатів; потік на огляд у середньому приблизно на 4,5% менший. Можливості першого огляду становили 280 еквівалентів повідомлень на тиждень: 35 годин роботи аналітиків по 8 еквівалентів повідомлень на годину. Ресурсу на рецензування не вистачає приблизно у 37% тижнів: у ці тижні надходження після вилучення дублікатів перевищують ресурс першого огляду в 280 еквівалентів повідомлень.

У цій записці поняття **придатне до дій** має визначення з інструкції служби: «Повідомлення є придатним до дій, якщо його огляд призводить щонайменше до одного з таких зафіксованих результатів: включення до щотижневого зведення про ситуацію; ескалації координатору служби для подальших дій; передачі партнерам спостереження щодо логістики чи доступу; або виправлення раніше поширеного твердження. Дублікати, неперевірені чутки, повідомлення поза тематикою та застарілі повідомлення не є придатними до дій».

2 Дані та позначки

Архів містить 18 240 публічних повідомлень приблизно з 30 відібраних стрічок: новин, офіційних і гуманітарних сторінок, а також відібраних аналітиками експортів із соціальних мереж. З них 1 679 (9,2%) були позначені як придатні до дій. Ще 819 майже дублікатів (4,5%) вилучено за допомогою нечіткого зіставлення заголовків: rapidfuzz token_set_ratio ≥ 88, той самий регіон і категорія, 72-годинне вікно.

Для розроблення моделі використано стратифікований аудиторський набір із 5 280 повідомлень, з яких 453 (8,6%) були придатними до дій. Набір поділено за часом, а не випадково, оскільки через конфлікт дані з часом змінюються.

Таблиця 1. Поділ аудиторського набору за часом

Частина	Тижні	Дати	Повідомлення	Придатні до дій	Частка придатних
Навчальна	0–38	3 березня – 30 листопада 2025	3 610	315	8,7%
Валідаційна	39–46	1 грудня 2025 – 25 січня 2026	910	77	8,5%
Тестова	47–64	26 січня – 31 травня 2026	760	61	8,0%
Аудиторський набір	**0–64**	**3 березня 2025 – 31 травня 2026**	**5 280**	**453**	**8,6%**

Відкладений тестовий набір (розділ 4) — це стратифікована аудиторська вибірка з 760 повідомлень за тижні 47–64, призначена для офлайнового оцінювання моделі. Проспективне тіньове тестування (розділ 5) охоплює весь потік надходжень (усі 1 842 повідомлення, що надійшли в тижні 58–64), і його результати подано окремо. Вибірка й тіньове тестування перетинаються в тижнях 58–64: повідомлення аудиторської вибірки за ці тижні входять до повного потоку тіньового тестування. Оскільки вони вимірюють різне (офлайнові показники ранжування проти проспективних операційних точності й повноти) і наведені в окремих розділах, жоден показник не враховано двічі.

3 Модель

Повідомлення ранжуються за допомогою логістичної регресії з L2-регуляризацією на 54 ознаках: 4 структурованих ознаки та 50 провідних термінів TF-IDF (відібраних за критерієм хі-квадрат лише на навчальних даних). Модель відкалібровано так, щоб її оцінки можна було читати як імовірності.

Чотири структуровані ознаки: надійність джерела; log(1 + кількість підтверджувальних повідомлень за останні 72 год); кількість термінових ключових слів; відстань до найближчого населеного пункту.

Базове правило оцінює кожне повідомлення за шкалою 0–100:

$$score = 100 \cdot \mathrm{clip}(0{,}45 \cdot credibility + 0{,}35 \cdot \min(\log(1 + corroboration_{72\mathrm{h}}) / 3;\ 1) + 0{,}20 \cdot category_prior;\ 0;\ 1)$$

Зважування класів	class_weight="balanced", що компенсує базовий рівень близько 9%.
Регуляризація	C = 3,0; обрано 5-кратною впорядкованою в часі перехресною перевіркою в межах навчального періоду (середнє CV AUC 0,8265).
Калібрування	Калібрування Платта (сигмоїдне), підігнане на валідаційному періоді (a = 0,768, b = −2,171).
Робочий діапазон	Повідомлення з каліброваною ймовірністю придатності до дій ≥ 0,16 (відображається як оцінка ≥ 16 за шкалою 0–100). В аудиторському наборі цей поріг відповідає 85-му процентилю, тобто охоплює приблизно 15% повідомлень із найвищими оцінками.

Оцінка базового правила за шкалою 0–100 — це окрема шкала для порівняння, яку використано лише в розділі 4; робочий діапазон визначено за каліброваною ймовірністю моделі.

4 Оцінювання моделі

Модель оцінювали на відкладеному тестовому наборі з 760 повідомлень, з яких 61 було придатним до дій. У таблиці 2 її порівняно з базовим правилом і з неінформативним орієнтиром.

Таблиця 2. Розрізнювальна здатність і калібрування на відкладеному тестовому наборі (760 повідомлень, 61 придатне до дій)

Показник	Модель	Базове правило	Без інформації
AUC (95% ДІ, метод DeLong)	0,801 (0,74–0,87)	0,746 (0,68–0,82)	0,5
Оцінка Браєра	0,0588	0,0646	0,0738
Нахил калібрування (1,0 = ідеал)	1,051	0,891	—

Примітка: позитивних випадків лише 61, тому інтервал AUC широкий, а інтервали двох моделей перетинаються.

5 Проспективне тіньове тестування

Тіньове тестування тривало впродовж тижнів 58–64 (13 квітня – 31 травня 2026) й охопило 1 842 повідомлення, з яких 164 (8,9%) були придатними до дій. Поріг робочого діапазону 0,16 перевищили 276 повідомлень (≈15% надходжень за цей період), і серед них було 79 із 164 придатних до дій повідомлень.

28,6%	3,2×	48,2%
Точність у робочому діапазоні	Від базового рівня 8,9%	Повнота

Позначки фіксували до будь-яких операційних дій, тому показник точності не завищено заднім числом.

В окремому дослідженні, описаному в дослідницькій записці «Прізми» *Вимірювання узгодженості щодо «придатності до дій»* (2025), на 200 повідомленнях, які незалежно позначили два аналітики, к Коена становив 0,634 — суттєва узгодженість за загальноприйнятими мірками. Навіть навчені аналітики лише частково погоджувалися щодо того, що є придатним до дій.

6 Прогнозування навантаження

Цільова змінна прогнозу — тижнева кількість повідомлень у робочому діапазоні (середнє 43,6, стандартне відхилення 16,5, пік 159).

Прогноз будується як ансамбль із рівними вагами моделі Голта із загасаючим трендом (без сезонної складової) та моделі ARIMA(1,1,1); обидві підігнано в NumPy пошуком по сітці за квадратами похибок прогнозу на один крок уперед. Оцінювали послідовним прогнозуванням із розширюваним вікном, що дало 17 прогнозів на тиждень уперед (тижні 48–64, 2 лютого – 31 травня 2026).

Таблиця 3. Точність прогнозів на тиждень уперед, тижні 48–64 (17 прогнозів)

Метод	sMAPE	MAE	Покриття P90
Ансамбль (модель Голта із загасаючим трендом + ARIMA(1,1,1))	16,1%	6,7	17/17
Наївний прогноз (минулий тиждень)	24,0%	9,8	16/17

Межі P90 розраховано за робастною оцінкою масштабу залишків (MAD 11,4 на тиждень).

7 Розподіл ресурсу на рецензування

Тижневий ресурс становить 48 годин роботи аналітиків: 35 годин першого огляду (280 еквівалентів повідомлень по 8 на годину) і 13 годин поглибленої роботи.

Ресурс розподіляється між 36 комірками (6 пріоритетних регіонів × 6 категорій). Кожна з 8 комірок із найвищою очікуваною цінністю отримує гарантований мінімум у 8 еквівалентів повідомлень — разом 64, або 23% ресурсу першого огляду. Показники черги та кадрового забезпечення розраховано на потоці після вилучення майже дублікатів.

Розподіл ресурсу перевірено на 2 000 сценаріях попиту Монте-Карло. Ці сценарії побудовано на основі 65 тижнів спостережуваних надходжень, і вони відтворюють історичну частоту тижнів, у які ресурсу не вистачає (приблизно 37%). Для кожного сценарію задача лінійного програмування (PuLP 2.7 + CBC) максимізує очікувану цінність огляду з урахуванням ресурсу та гарантованих мінімумів. Ресурсу не вистачає в 744 сценаріях (37%), а гостро не вистачає — у 63 (3%).

Таблиця 4. Порівняння методів розподілу на 2 000 сценаріях попиту Монте-Карло

Метод	Покриття цінності огляду			Покриття критичних комірок, P10	Недопокриття найвищого пріоритету
	Усі сценарії	Ресурс обмежений	Гострий дефіцит		
Рівні години	91%	88%	71%	82%	13,1%
Жадібний за оцінкою	99%	98%	81%	99%	0,8%
Жадібний за цінністю, зваженою на терміновість	99%	98%	82%	99%	0,6%
LP Монте-Карло	99%	98%	82%	99%	0,6%

Ресурс обмежений: 744 сценарії (37%), у яких ресурсу не вистачає. Гострий дефіцит: 63 сценарії (3%), у яких ресурсу гостро не вистачає.

Перевірка на піковому тижні (942 повідомлення надійшло; 920 після вилучення дублікатів): значення цільової функції LP = значення цільової функції жадібного методу = 74,1.

8 Операційний ефект (моделювання)

Щоб оцінити операційний ефект ранжованого огляду, ми провели моделювання черги на всіх 455 днях архіву за пропускної здатності 40 еквівалентів повідомлень на день. Наведені нижче результати отримано моделюванням; це не вимірювання в реальній роботі.

Таблиця 5. Змодельована своєчасність огляду: порядок надходження (FIFO) проти ранжованого порядку Моделювання

Показник	FIFO	Ранжований
Медіанний час до першого огляду, усі повідомлення	57 год	3 год
Медіанний час до першого огляду, придатні до дій	56 год	1 год
Придатні до дій, переглянуті протягом 24 год	33%	95%
Найвищий пріоритет, переглянутий протягом 48 год (піковий тиждень)	11%	100%

Моделювання черги на всіх 455 днях архіву за 40 еквівалентів повідомлень на день. Не є вимірюванням у реальній роботі.

9 Інструменти

Фінальний аналіз	Python; NumPy (ETS та ARIMA, підігнані безпосередньо); scikit-learn (TF-IDF, відбір термінів за хі-квадрат, логістична регресія, впорядкована в часі перехресна перевірка, сигмоїдне калібрування); SciPy (довірчі інтервали AUC за методом DeLong); rapidfuzz (виявлення дублікатів); PuLP 2.7 + CBC (LP для планування кадрового забезпечення).
Робоча система	feedparser, BeautifulSoup, pandas (отримання й розбір даних); DuckDB (сховище повідомлень); Streamlit (панель моніторингу, запускається локально); Git, GitHub Actions, pytest (контроль версій і тестування).

10 Ролі

Проєкт виконувала служба моніторингу впродовж 15 місяців. Координація служби охоплювала дизайн дослідження, вирішення спірних позначок, ескалації та контроль якості. Кількісні компоненти (модель ранжування, прогнозувальник навантаження, розподільник ресурсу на рецензування) розроблено у співпраці зі службою. Аналітики-волонтери займалися маркуванням повідомлень, добором джерел і експортами із соціальних мереж. Управління даними охоплювало стрічки, перевірки на дублікати та зберігання даних.

11 Обмеження

- Тестовий набір містить лише 61 придатне до дій повідомлення, тому інтервал AUC широкий.
- Міжекспертна узгодженість к = 0,634 обмежує, наскільки добре будь-яка модель може оцінювати повідомлення.
- Дані змінюються впродовж 65 тижнів конфлікту; поділ за часом захищає від витоку даних, але не від подальшого дрейфу.
- Через лише 17 тестових тижнів оцінки sMAPE та покриття P90 неточні.
- Остаточні показники отримано повторним запуском конвеєра на повному архіві, тож робоча система могла в деталях відрізнятися.
- Скорочення часу до огляду отримано з моделювання, а не з реального чи рандомізованого порівняння.
- Система ранжує та рекомендує; усі рішення ухвалювали люди, і записка не робить причинно-наслідкових висновків.

Дослідницький центр «Прізма» · research@prismresearch.org.ua · prismresearch.org.ua