

News Impact

агрегатор найчастіше згаданих новин

Кеш → Переклад → Ембеддинги → Кластеризація → Дедуп eTLD+1 → Top-N → Підсумки

Проблема і мотивація

- Інфошум із сотень заголовків різних ЗМІ.
- Потрібен автоматичний огляд **спільних сюжетів** і стислий/детальний підсумок.

Цілі та функціональність

- Виявлення спільних сюжетів, Top-N за унікальними виданнями (eTLD+1).
- Підсумки для будь-якої теми, включно зі складними умовами.
- Кеш заголовків/тіл/підсумків; вибір мови відображення (EN/UA).

Архітектура (огляд)

```
RSS → Ingest → Titles cache
      ↳ Bodies → Articles cache
Translate (idempotent) → Embeddings → Clustering → Stories
Stories → Dedup eTLD+1 → Rank Top-N → LLM summary
Topic(tab) → LLM select → Deep summary (EN) → optional UA → save .txt
```

Деталі реалізації

- `feedparser` + `requests` + BeautifulSoup (AMP-fallback).
- Переклад заголовків/тіл: OpenAI, sha1-кеш.
- Ембеддинги: `all-MiniLM-L6-v2` .
- Кластери: косинус, жадібно; поріг у UI.
- Ранжування: унікальні видання → згадки → рання поява.

Controls

RSS feeds

feeds.bbc.co.uk



cnn.com



feeds.finance.yahoo.com



investing.com



Max items (ingest)

1600



Max age (hours)

48



Similarity threshold

0.72

Top-N stories

10

OpenAI model

gpt-4.1-mini



☒ Translate titles to English

☐ Translate article bodies (costly)

Article timeout (s)

10



Update cache

Summarize

News Impact — Top Most-Mentioned Stories (distinct outlets)

Top stories

All titles

Top 10 most-mentioned stories (by distinct outlets)

Bail has been posted for MP Kuznetsov, who is suspected of corruption in the procurement of drones and electronic warfare equipment.

Bail has been posted for MP Oleksiy Kuznetsov, who is suspected of corruption related to the procurement of drones and electronic warfare equipment. The High Anti-Corruption Court set his bail at eight million hryvnias, lower than the prosecution's demand of 30 million. Kuznetsov, a member of the Servant of the People party, was among four individuals detained in an investigation revealing a scheme involving up to 30% kickbacks in state contracts.

outlets: 4 · first seen: 2025-09-02T02:47:58.301155+00:00

- [CENSOR.NET] [Corruption in drone procurement: an 8 million hryvnia bail was posted for "Servant of the People" member Kuznetsov](#)
- [NV.UA] [Bail has been posted for MP Kuznetsov, who is suspected of corruption in the procurement of drones and electronic warfare equipment.](#)
- [PRAVDY.COM.UA] [Corruption in the procurement of electronic warfare systems and drones: a bail of 8 million was posted for MP Kuznetsov](#)
- [TSN.UA] [Corruption in the procurement of drones and electronic warfare systems: bail posted for MP Kuznetsov](#)

Ukrainian soldiers have liberated Novoekonomichne in the Donetsk region from Russian forces (video).

Ukrainian forces have liberated the village of Novoekonomichne in the Donetsk region from Russian control, as reported by the General Staff of the Armed Forces of Ukraine. The operation, carried out by the 425th Separate Assault Regiment "Skelya," took about two weeks and involved raising the Ukrainian flag in the settlement. This action is part of ongoing efforts by Ukrainian troops to regain control over temporarily occupied areas in Donetsk.

outlets: 4 · first seen: 2025-09-02T03:11:15.512902+00:00

- [CENSOR.NET] [The Armed Forces of Ukraine have liberated the village of Novoekonomichne in the Donetsk region and raised the Ukrainian flag there, according to the General Staff. VIDEO](#)
- [RBC.UA] [Ukrainian soldiers have liberated Novoekonomichne in the Donetsk region from Russian forces \(video\).](#)
- [TSN.UA] [In Donetsk region, the Armed Forces of Ukraine have liberated the settlement of Novoekonomichne from the occupiers.](#)
- [UKRINFORM.NET] [Ukrainian forces liberate Novoekonomichne in Donetsk region](#)

Ukraine has likely struck a Russian target with the "Fleming" missile for the first time. What is known about this?

Ukraine likely conducted its first strike using the domestically produced "Fleming" cruise missile on a Russian target in annexed Crimea near Armyansk on August 30. The missile reportedly has a range of up to 3,000 kilometers and can carry a warhead weighing over a ton, potentially allowing Ukraine to target deep into Russian territory. While video footage and satellite images suggest damage to a Russian FSB border post, independent confirmation and detailed damage assessments remain unavailable.

outlets: 4 · first seen: 2025-09-02T03:11:15.512902+00:00

Тема за запитом

- Відбір релевантних заголовків з усіх перекладених.
- Узагальнення повних текстів у детальний EN-підсумок ($\approx 300\text{--}500$ слів).
- Перемикач EN/UA: переклад **після** генерації, збереження в `.uk.txt`.
- Повторне використання файлів за датою.

Метрики та оцінка

- `items_ingested` , кластери, унікальні видання в Top-N.
- Частка `body_len>0` , наявність `*_en` .
- Кількість нових LLM-викликів за сесію.

Запуск

```
mamba env create -f requirements.yml  
mamba activate news-impact  
streamlit run streamline.py
```

Висновки

Система надійно агрегує, кластеризує та узагальнює новини, надаючи цінні дайджести для **будь-яких тем** — від війни до технологій та економіки — із прозорим кешуванням і відтворюваними файлами результатів.