

# AI, Security & Defense Forum (AISD)

Where AI/ML meets Security & Defense



Universitat  
Pompeu Fabra  
Barcelona



AISD



Fall 2025 Series on “Influence Operations”

## ***Characterizing Knowledge Manipulation in a Russian Wikipedia Fork***

by **Mykola Trokhymovych**

(Universitat Pompeu Fabra, Spain)

**Time:** 23 October 2025 12h00 Boston (18h00 Oslo)

**Hosts:** Phil Tinn & Olivia Houck





# About me:



# Mykola Trokhymovych

[trokhymovych.com](https://trokhymovych.com) | [mykola.trokhymovych@upf.edu](mailto:mykola.trokhymovych@upf.edu)

- PhD Candidate @ Universitat Pompeu Fabra (UPF)
- Visiting Researcher, Indiana University (OSoMe)
- Trained as a Data Scientist and Systems Analyst
- Over 5 years of researching the Wikimedia ecosystem

## Conferences:

- ACL'25, ICWSM'25, ACL'24, KDD'23, CIKM'21

**Why Wikipedia?**

**RWFork background**




**Collecting dataset of changes**

**Change analysis and insights**

**Why is it important? & What's next?**

**Q&A**

# Why Wikipedia?

-  **A Leading Knowledge Source on the Web**
  - Ranks **6th** among the most important sources of historical information, ahead of museum visits, university courses, and social media (Burkholder & Schaffer, 2021).
  - Available in **300+** languages, with **60M+** articles on Wikipedia.
-  **Transparency and Open Source**
  - Created and maintained by a global community with the goal of becoming “the sum of all human knowledge” (Sutcliffe, 2016).
  - Wikipedia exists thanks to voluntary donations from millions of users worldwide.
  - Operates under clear community rules, helping to prevent hidden interests and bias.
-  **Impact Beyond the Wikimedia Ecosystem**
  - Appears in results from Google, ChatGPT, Siri, Alexa, and others.
  - Used as a training source for artificial intelligence (AI), including LLMs.

# Wikipedia and Censorship

Some actors disagree with Wikipedia's content and policies.

## Blocking Wikipedia:

Online encyclopedia Wikipedia blocked in China ahead of Tiananmen anniversary

By Josh Horwitz

May 15, 2019 12:02 PM GMT+2 · Updated May 15, 2019



Turkey blocks Wikipedia under law designed to protect national security

Users trying to access online encyclopaedia via Turkish internet providers receive 'connection timed out' error message



## Creating an Alternative:

Conservapedia - the US religious right's answer to Wikipedia



How Baidu built an encyclopedia with 16 times more Chinese entries than Wikipedia

Transition from hard format encyclopaedias to open-source encyclopaedias has raised some quality-of-information issues though



**Hamichlol** (Hebrew: המכלול "The Entirety") is a **censored wiki-based Internet encyclopedia project** for the **Haredi public**. Most of the articles on the site are exact copies (as of the date they were copied) of articles from the **Hebrew Wikipedia** or re-edited versions of Hebrew Wikipedia articles, adapted for the Haredi reader, while



# Internet Governance in Russia

The early development of the Internet in Russia occurred without state control, resulting in a complex and decentralized network (Glasze et al., 2023).

Roskomnadzor has to rely on the voluntary cooperation of autonomous systems and providers, but a lack of trust reduces its effectiveness (Ermoshina & Musiani, 2022).

Major Internet companies (Yandex, VK, Mail.ru) are closely linked to the state and act as intermediaries of censorship (Zuckerman, 2010). However, the government continues to seek new forms of online control (Wijermars, 2021).



**Y**andex



# Creating an Alternative Wikipedia (RWFork)

In July 2023, Ruwiki (RWFork) was launched as a fork of the Russian Wikipedia, created to “comply with Russian legal requirements.”

The new project is led by a long-time Russian Wikipedia top editor.

This followed the Russian government imposing fines on Wikimedia totaling around \$255 million.

## In Putin’s Russia, ‘fake news’ now means real news

A term coined by misinformation researchers was coopted by President Trump. It has become a tool of authoritarians.

## Russia fines Wikimedia Foundation over Ukraine war entries

By Reuters

≡ Bloomberg

Subscribe



Businessweek | Technology

## Russian Wikipedia’s Top Editor Leaves to Launch a Putin-Friendly Clone

The move is seen as a potential precursor to a ban on the original version.



TIME

SUBSCRIBE

WORLD • RUSSIA

## Russia to Create Its Own ‘Alternative Wikipedia’

# How has Wikipedia's knowledge been manipulated in RW/Fork?

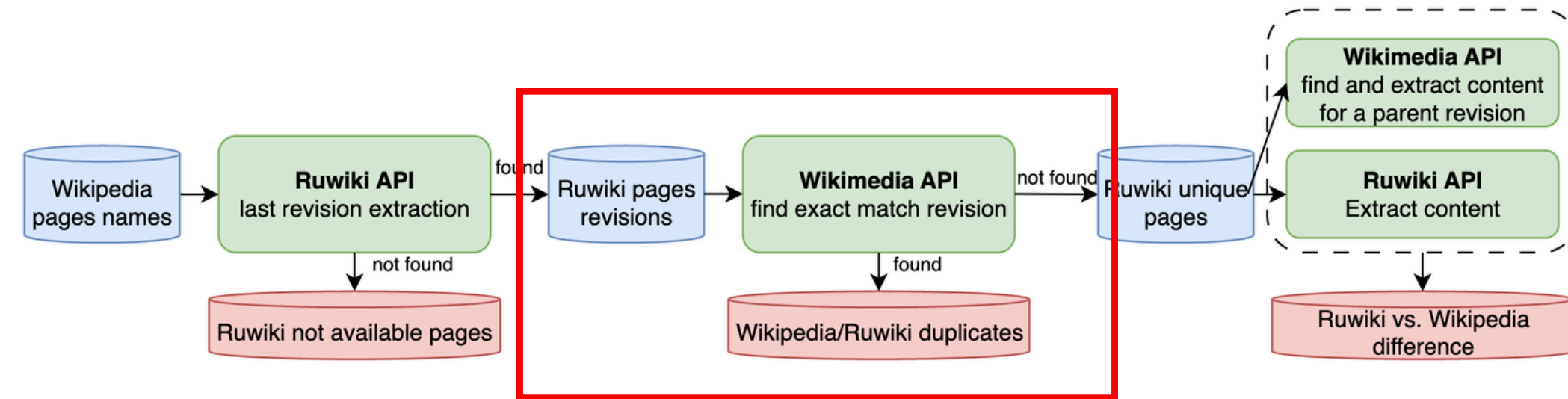
**33,664** (1.75%)

pages changed in first 3 month





# Data collection



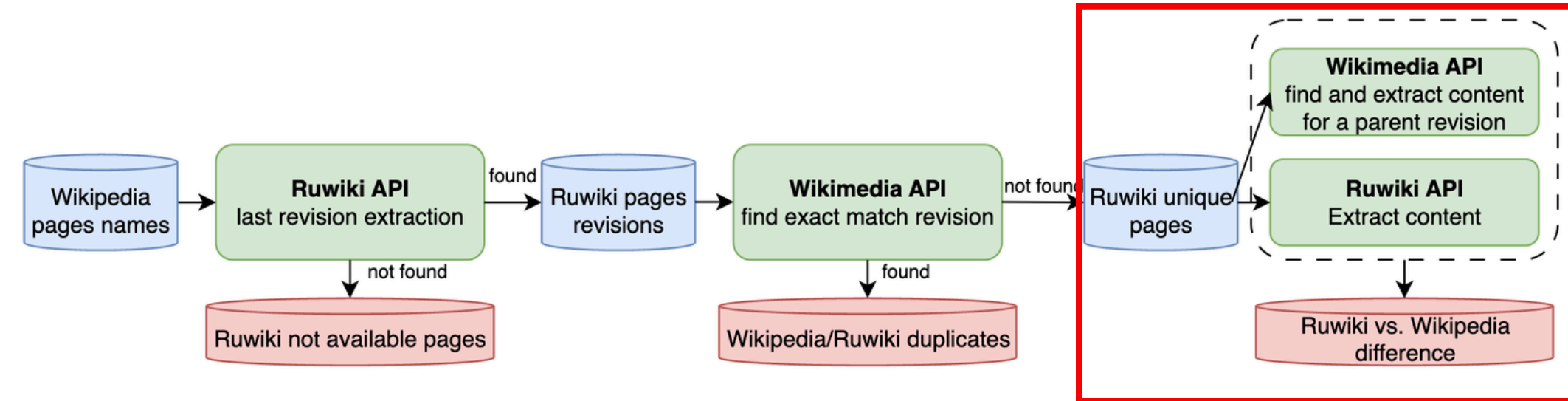
## Step 2. Defining the initial state of the page

- ◆ Compared the revision histories between Russian Wikipedia and RWFork.
- ◆ Identified the “parent” revision from Russian Wikipedia – the page version that served as the starting point for RWFork.

✓ (текущ.   пред.)	○	16:40, 20 июня 2023	Sanopus (обсуждение   вклад) .. (120 278 байт) (-559) .. (оформление) (отменить) [отпатрулирована автоматически]
✓ (текущ.   пред.)	○	07:13, 19 июня 2023	Sunshine (обсуждение   вклад) .. (120 837 байт) (+1296) .. (отменить) [отпатрулирована автоматически]
✓ (текущ.   пред.)	○	21:35, 17 июня 2023	Sanopus (обсуждение   вклад) .. (119 541 байт) (-17 258) .. (rv) (Метка: отмена) [отпатрулирована автоматически]
(текущ.   пред.)	○	21:29, 17 июня 2023	(имя автора стёрто) .. (136 799 байт) (-2) .. (описание правки стёрто) (Метка: отменено)
(текущ.   пред.)	○	21:26, 17 июня 2023	(имя автора стёрто) .. (136 801 байт) (+17 262) .. (описание правки стёрто) (Метки: ручная отмена, отменено)
(текущ.   пред.)	○	21:23, 17 июня 2023	(имя автора стёрто) .. (119 539 байт) (-2) .. (описание правки стёрто) (Метка: отменено)
(текущ.   пред.)	○	19:03, 7 июня 2023	Sanopus (обсуждение   вклад) .. (119 541 байт) (-95) .. (оформление) (отменить) (Метка: отменено)
(текущ.   пред.)	○	18:19, 7 июня 2023	Sanopus (обсуждение   вклад) .. (119 636 байт) (-145) .. (оформление) (отменить) (Метка: отменено)
(текущ.   пред.)	○	16:36, 6 июня 2023	Sanopus (обсуждение   вклад) .. (119 781 байт) (-17 020) .. (707437-НВ) (отменить) (Метки: отменено, редактор вики-текста 2017)
✓ (текущ.   пред.)	○	18:16, 2 апреля 2023	© Перенесено из ru.wikipedia.org/wiki/Херсон по лицензии CC BY-SA 3.0. Список авторов источника в «Истории». В страницу вносились изменения после переноса. (Unknown user) .. (136 801 байт) (0) .. (Изменил настройки защиты для «Херсон»: Автоматическая полузащита статьи из списка потенциально опасных для редакторов ([[Редактирование=только автоподтверждённые]] (бессрочно) [[Переименование=только автоподтверждённые]] (бессрочно))) (отменить) [отпатрулирована автоматически]
✓ (текущ.   пред.)	○	09:48, 23 марта 2023	© Перенесено из ru.wikipedia.org/wiki/Херсон по лицензии CC BY-SA 3.0. Список авторов источника в «Истории». В страницу вносились изменения после переноса. (Unknown user) .. (136 801 байт) (-21) .. (→Ссылки) (отменить) [отпатрулирована автоматически]
✓ (текущ.   пред.)	○	13:32, 7 февраля 2023	© Перенесено из ru.wikipedia.org/wiki/Херсон по лицензии CC BY-SA 3.0. Список авторов источника в «Истории». В страницу вносились изменения после переноса. (Unknown user) .. (136 822 байта) (-30) .. (→Фотогалерея Херсона) (отменить) [отпатрулирована автоматически]
✓ (текущ.   пред.)	○	04:35, 4 февраля 2023	© Перенесено из ru.wikipedia.org/wiki/Херсон по лицензии CC BY-SA 3.0. Список авторов источника в «Истории». В страницу вносились изменения после переноса. (Unknown user) .. (136 852 байта) (-2324) .. (→СМИ: по шаблону) (отменить) (Метка: правка через визуальный редактор) [отпатрулирована автоматически]

Revision history of the RWFork page.  
The parent revision (the last one created in Wikipedia) is highlighted.

# Data collection



## Step 3. Extracting the changes in content

◆ Detected and processed text changes using the procedure from previous work, which is based on the mwedittypes tool.

### Laurit Krasniqi: Difference between revisions

From Wikipedia, the free encyclopedia

Browse history into

Revision as of 11:13, 29 July 2022 (edit)  
ManiacOfSport (talk | contribs)  
← Previous edit

#### Parsed text changes from wikitext:

(1) text inserts: [], (2) texts removed: []  
(3) text changes: [('Born in Belgium, he has chosen to represent the Kosovo national team.', 'Born in Kosovo, he has chosen to represent the Kosovo national team.')]

Line 21:

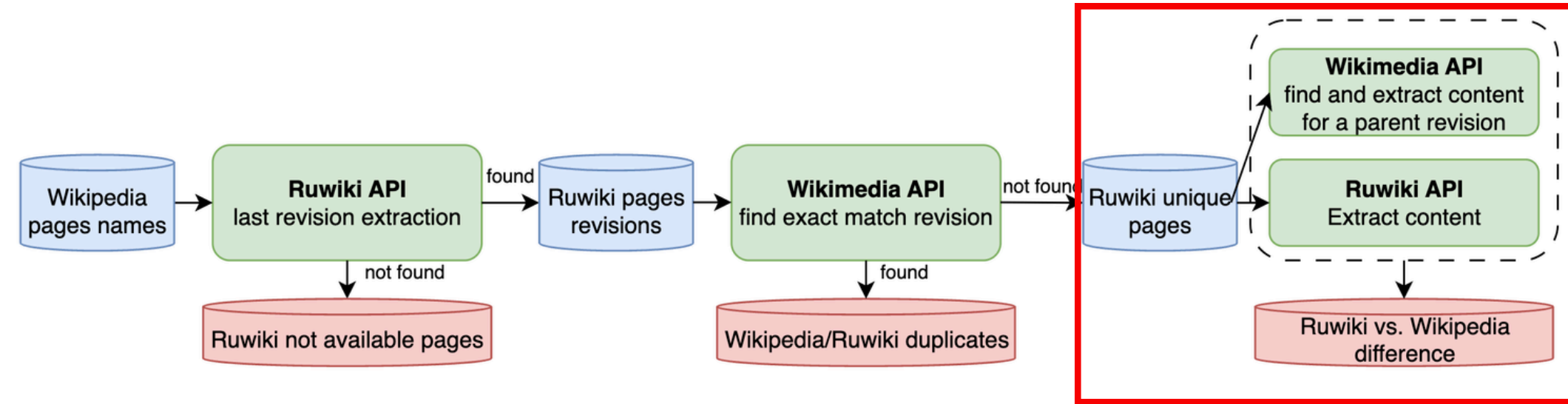
```
|club-update = 22 May 2022
}}
'''Laurit Krasniqi''' (born 14 July 2001) is a professional [[Association football|footballer]] who plays as a [[midfielder]] for Belgian club [[Royal Antwerp F.C.|Antwerp]]. Born in [[Belgium]], he has chosen to represent the [[Kosovo national football team|Kosovo national team]].
```

Line 21:

```
|club-update = 22 May 2022
}}
'''Laurit Krasniqi''' (born 14 July 2001) is a professional [[Association football|footballer]] who plays as a [[midfielder]] for Belgian club [[Royal Antwerp F.C.|Antwerp]]. Born in [[Kosovo]], he has chosen to represent the [[Kosovo national football team|Kosovo national team]].
```

The process of extracting the content change between two page versions. Based on Trokhymovych et al. (2023). *Fair Multilingual Vandalism Detection for Wikipedia*. (KDD '23)

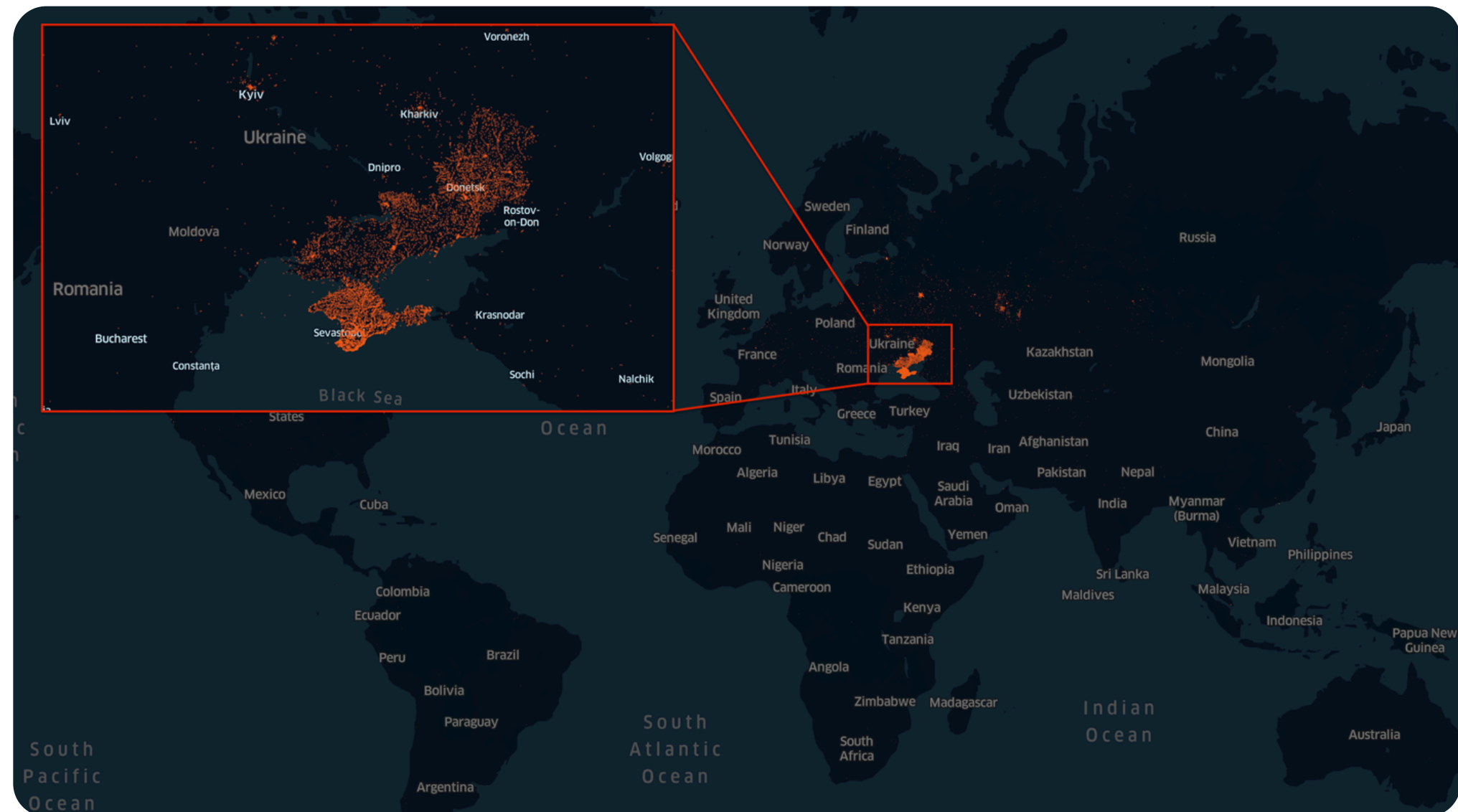
# Data collection



## Step 4. Extending dataset page features

◆ Enhancing the dataset by extracting:

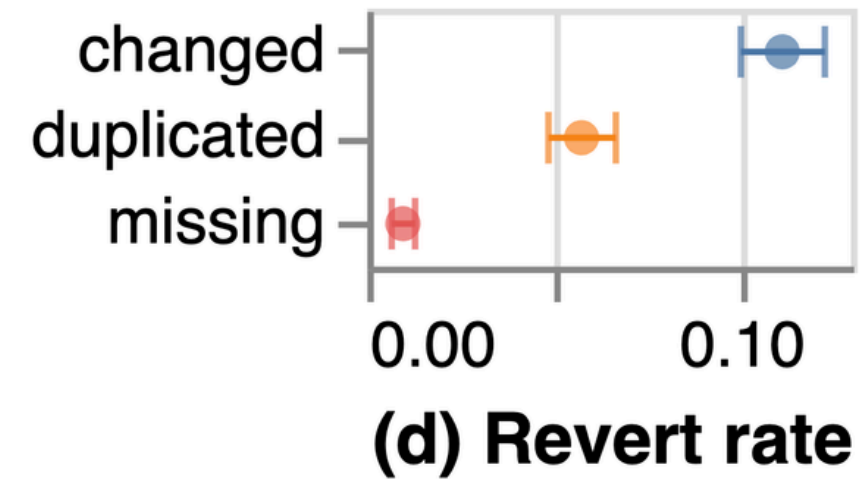
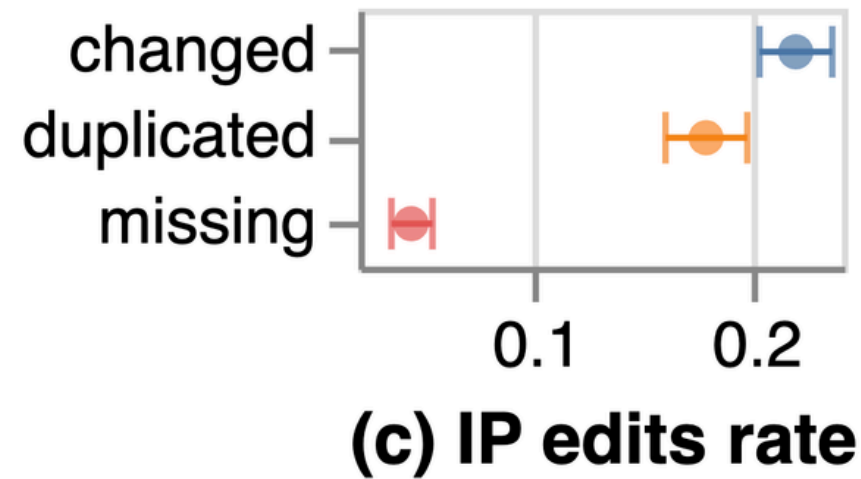
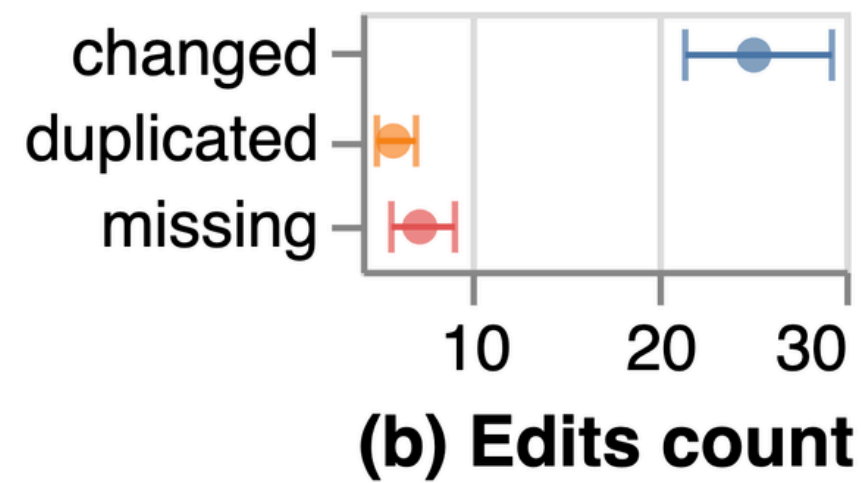
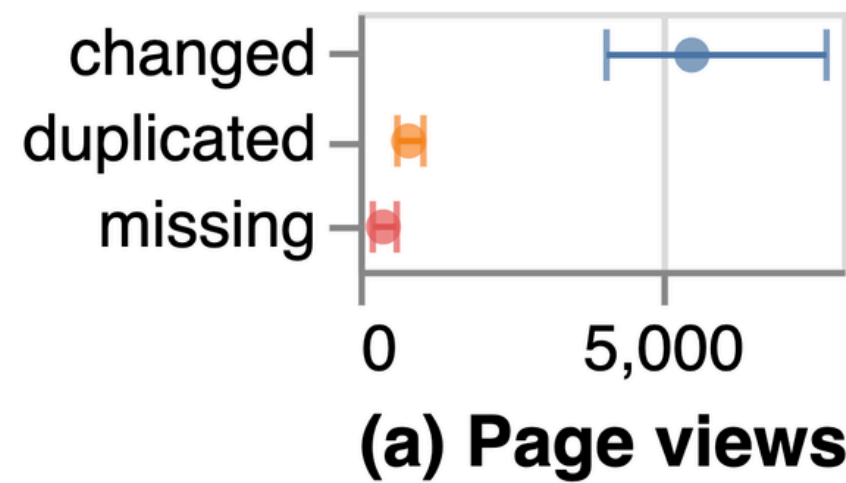
- changed references
- pages location
- page topics
- engagement metrics
- etc...



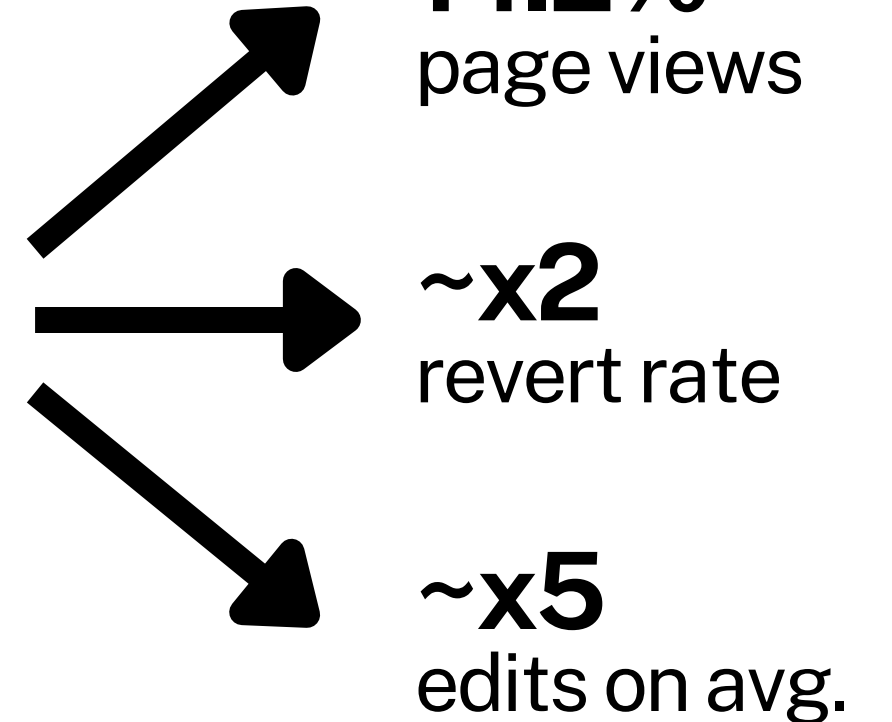
Locations of pages changed in RWForK

# Initial Data Analysis: Key Insights

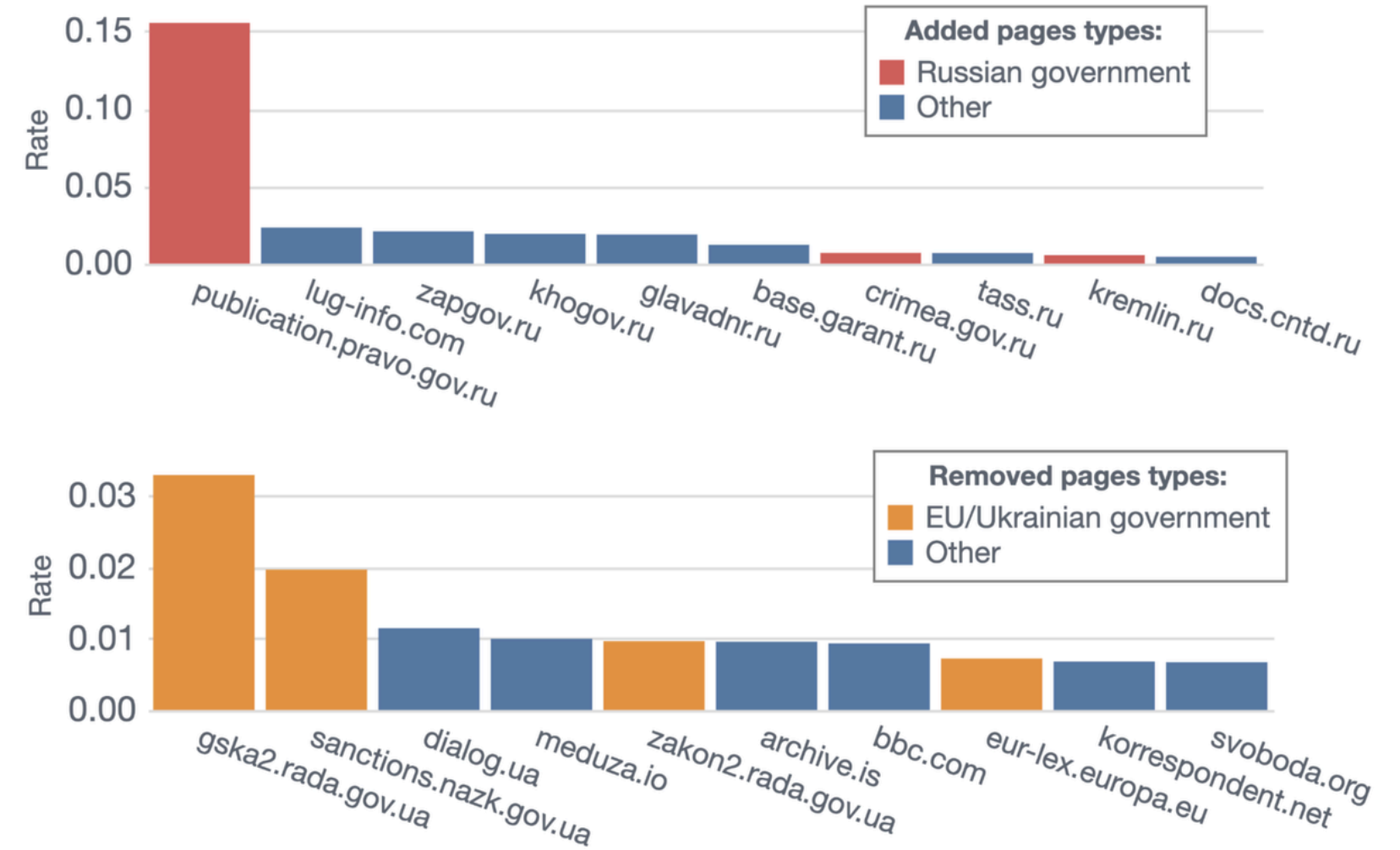
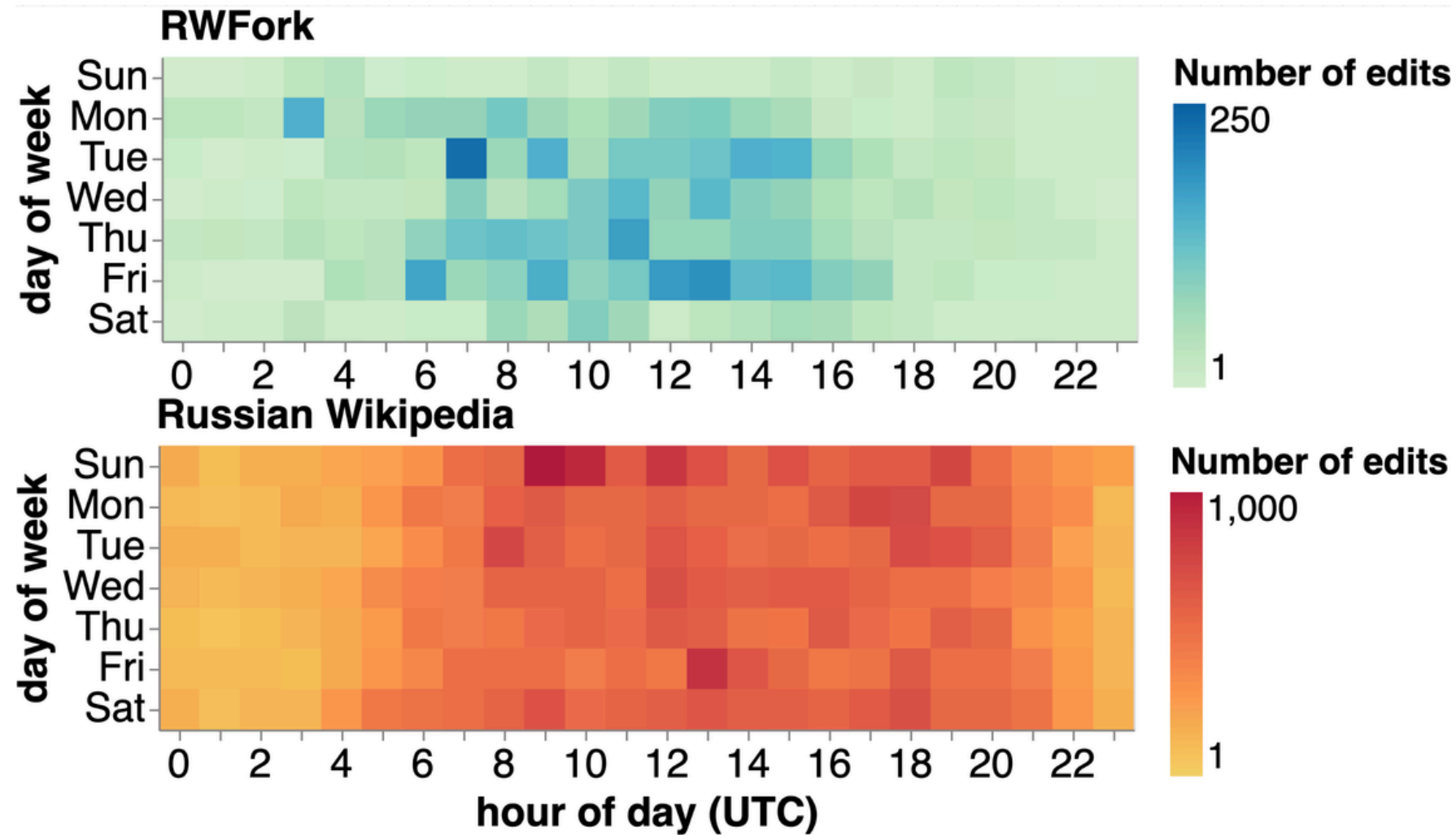
# Changing influential and controversial pages



**1.75%**  
articles changed



# Coordinated Editing Patterns



- Activity peaks during office hours on weekdays
- Adds Russian Government sources, removes EU and Ukrainian
- Focus on pages related to Ukraine, Russia, and Belarus

# Content changes examples

Page title	Content changes
Zburevsky Kut	<b>Lines changed:</b> [ ("Zburevsky Kut or Zburevsky Liman is a lake, bay in the Dnipro delta, located on the territory of the Skadovsky district ( <i>Kherson region, Ukraine</i> ) "Zburevsky Kut or Zburevsky Liman is a lake, a bay in the Dnipro delta, located on the territory of the Skadovsky district <i>Kherson region of Russia</i> ")]
Medal "For the Return of Crimea"	<b>Categories removed:</b> ["Articles with edit notes about caution when editing"]
Vladimir Samokish	<b>Lines changed:</b> [("Due to Russia's invasion of Ukraine, he is under international sanctions of the European Union, USA, Great Britain, and other countries", "He is <i>on the sanctions list</i> of the European Union, USA, Great Britain, and other countries")] <b>Categories changes:</b> ["Persons under the sanctions related to the conflict in Ukraine"]
Trimetozine	<b>Tag added:</b> Legislation of the Russian Federation—Medicines
Kendra (name)	<b>Lines deleted:</b> ["1981) - American porn actress and erotic model", "Sunderland, Kendra (born — 1995) - American porn actress and erotic model"]
Svoya igra	<b>Lines changed:</b> [("Broadcast on NTV on <i>Sundays</i> at 15:00", "Broadcast on NTV on <i>weekends</i> at 15:00")]

Random sample of content changes of various types made by RWFork editors (original Russian texts translated to English).

# **Systematization of Content Changes**

# Systematization of Content Changes

## Step 1. Summarization (characterisation) of edits:

*Text to summarize example:*

**page title:** Калиновское;

**lines changed:**

*Old value:* Калиновское — село в Белокуракинском районе Луганской области

*New value:* Калиновское — село в Белокуракинском районе Луганской Народной Республики

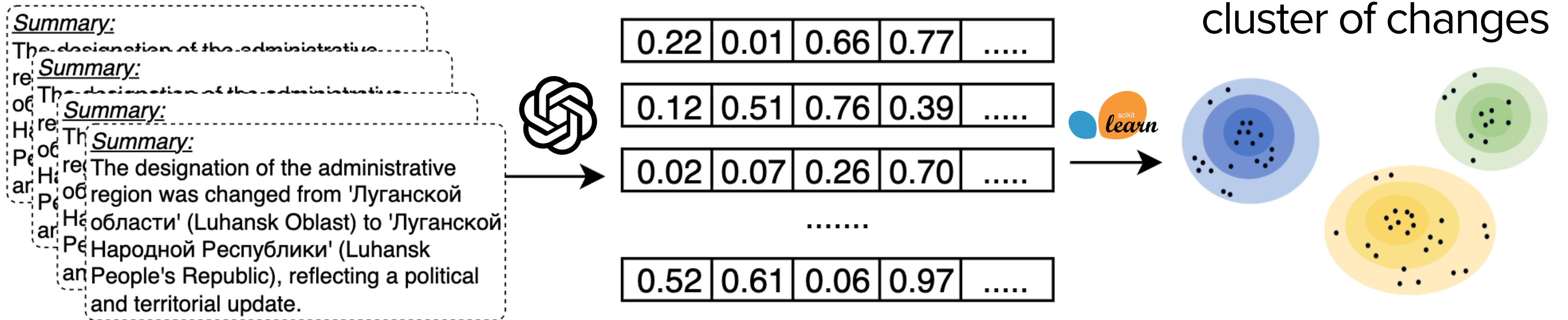


*Summary:*

The designation of the administrative region was changed from 'Луганской области' (Luhansk Oblast) to 'Луганской Народной Республики' (Luhansk People's Republic), reflecting a political and territorial update.

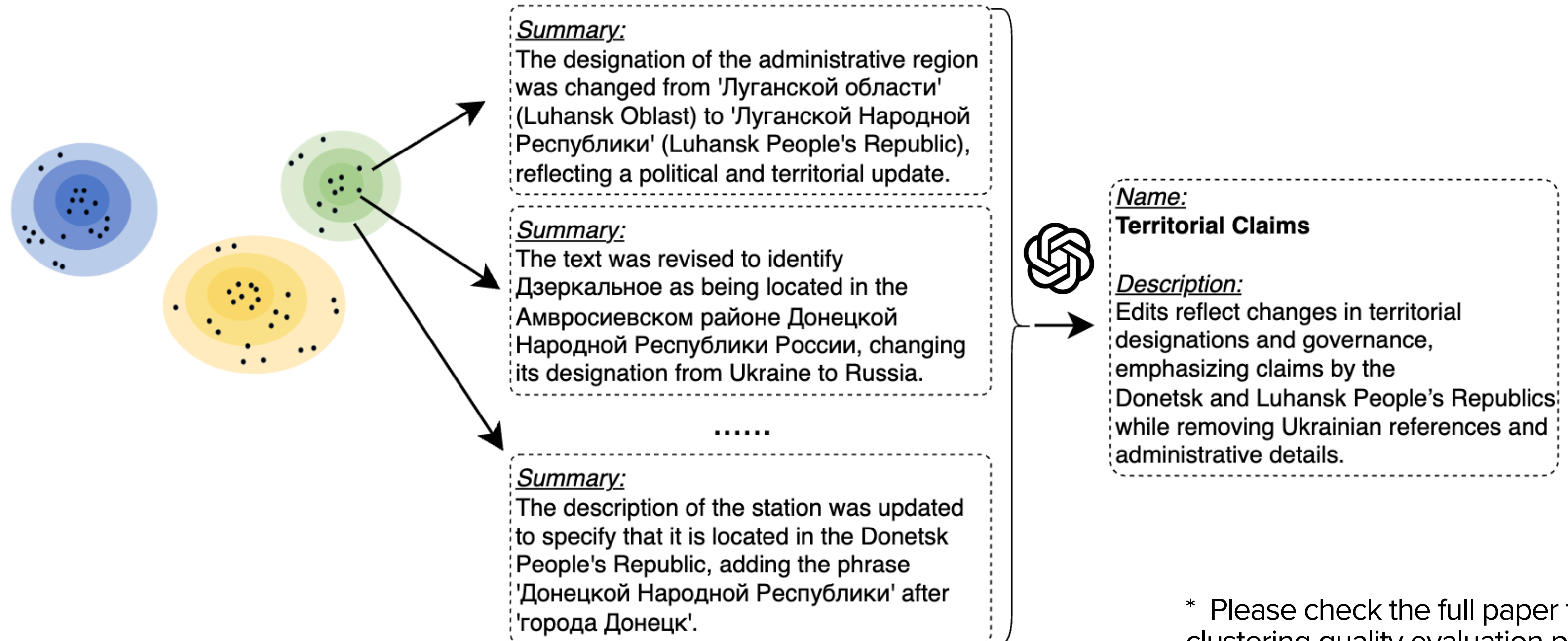
# Systematization of Content Changes

## Step 2. Getting vector representation and clustering:



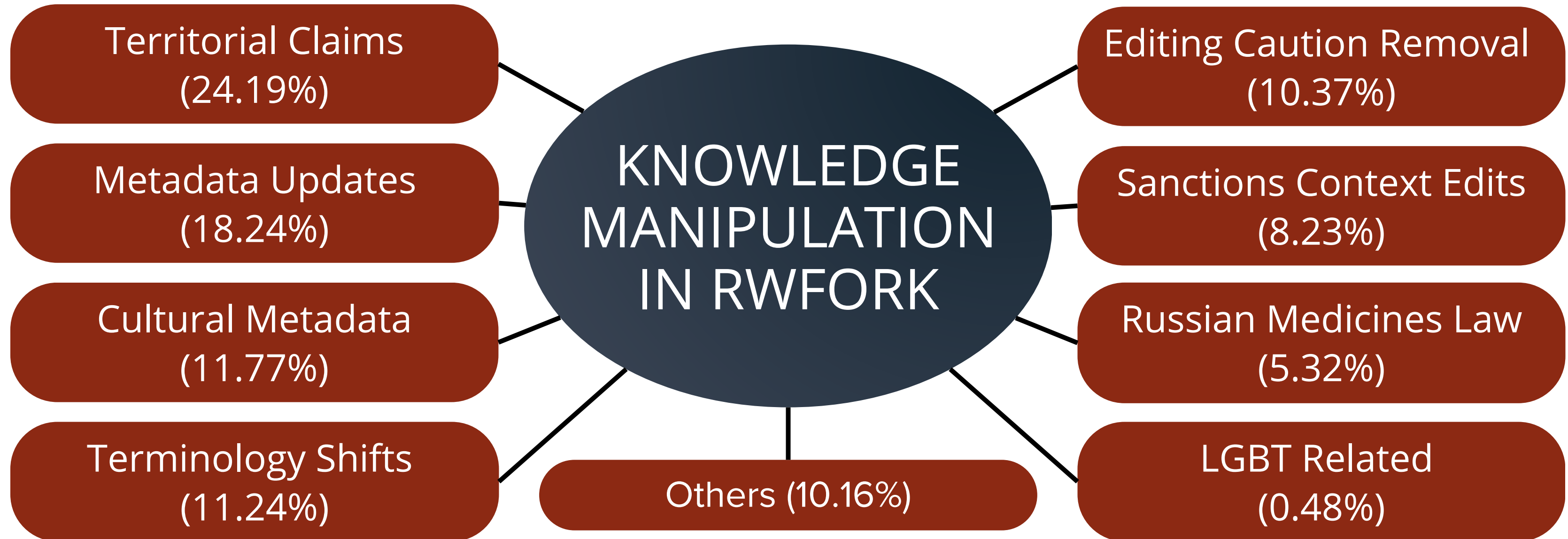
# Systematization of Content Changes

## Step 3. Building the taxonomy of changes\*

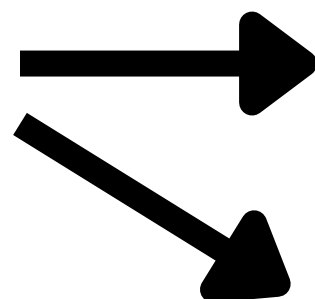


\* Please check the full paper for the clustering quality evaluation procedure

# Systematic Knowledge Manipulation



~90%  
changes



8  
categories

~44%

related to the Russo-Ukrainian War

# Summary

First study of how original Wikipedia content has been **forked** and **manipulated** to meet the requirements of a national regulation.

- RWFork edits target **top visited and controversial pages**
- Editors follow **office-hour activity patterns**
- Changes are **systematic**, mostly focused on the Russo-Ukrainian War

RWFork dataset is available for further study:  
<https://zenodo.org/records/15073728>



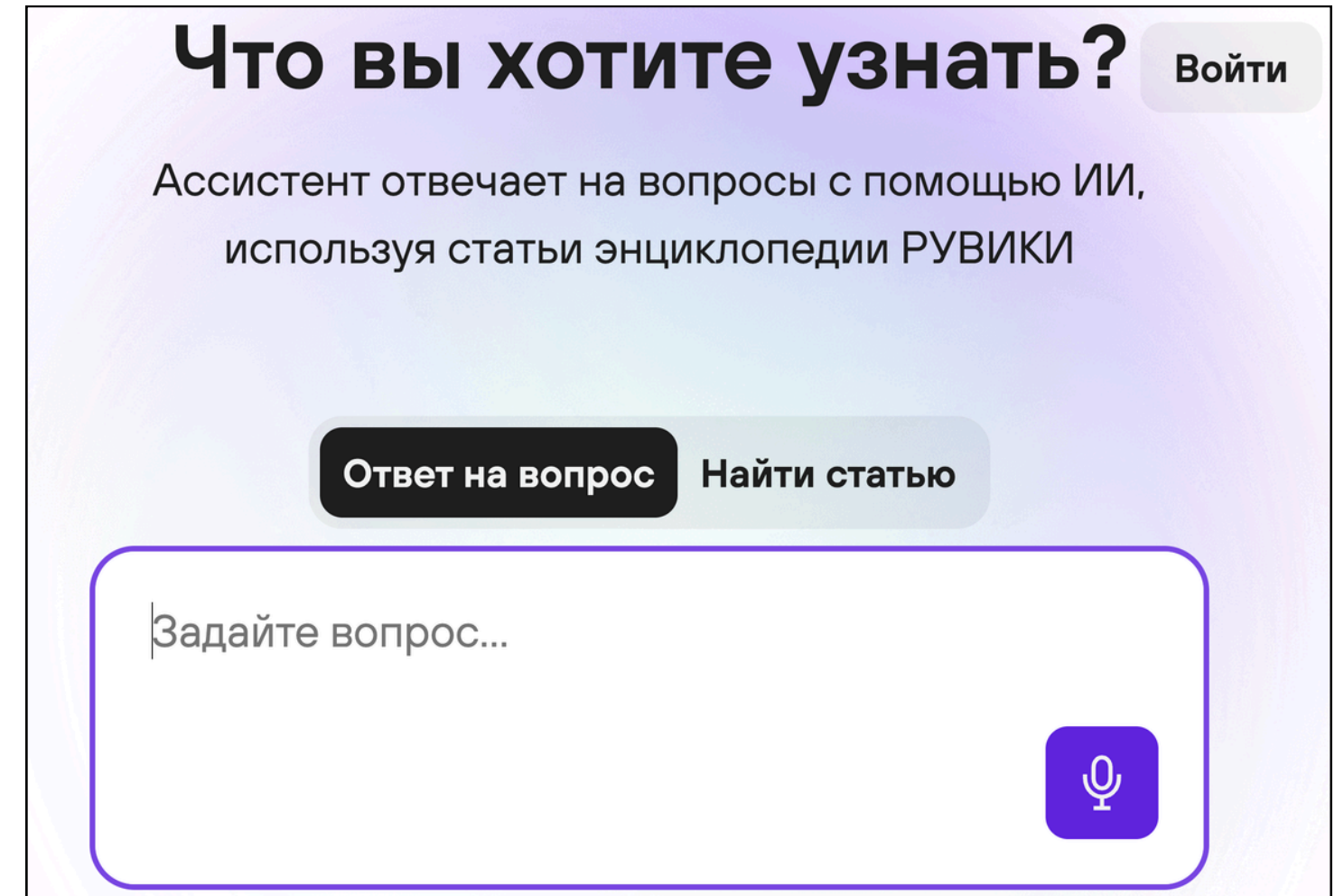
# Why is it important?

## Short-term:

- Amplification & institutionalization of specific websites (references).
- Risks of using data in training and inference for modern AI models.

## Long-term:

- Legitimize certain ideologies in law, education, and media regulation.
- It can influence national identity, historical memory, and international perception.



ruwiki.ru statistics based on Similarweb (September 2025)

# Why is it important?

- AI models and search-augmented chatbots treat encyclopedias as trusted sources.
- RWFork (or any other alt-wikipedia) can act as a shadow Wikipedia, targeting AI crawlers rather than human readers.
- What is the next target?

**The Washington Post**  
*Democracy Dies in Darkness*

**Russia seeds chatbots with lies. Any bad actor could game AI the same way.**

In their race to push out new versions with more capability, AI companies leave users vulnerable to “LLM grooming” efforts that promote bogus information.

April 17, 2025



# Ruwiki chatbot answers\*

## What were the Russian army's losses during the invasion of Ukraine?



It's possible they were referring to losses in the Ukrainian Armed Forces (UAF). As of August 2025, according to official data from the Russian Ministry of Defense, Ukraine had lost **over 1.7 million** service members killed and wounded during the military operations.

## Which Orthodox Church is correct in Ukraine?



Perhaps they meant **the Ukrainian Orthodox Church (UOC)**, which until 2019 was recognized by all local Orthodox churches as the only canonically legitimate Orthodox Church in Ukraine (as part of the Moscow Patriarchate).

However, on August 20, 2024, the Verkhovna Rada adopted a law banning the canonical Ukrainian Orthodox Church. On August 17 of the same year, the All-Ukrainian Council of Churches and Religious Organizations supported the bill banning the Ukrainian Orthodox Church.

## Was the annexation of Crimea legal?



**The annexation of Crimea to the Russian Federation in 2014 was recognized as a legal decision .**

## Why is Russia fighting "LGBT propaganda"?

Russia is cracking down on "LGBT propaganda" **amid concerns about the influence of non-traditional Western LGBT culture on children and young people .**

## Is it true that the Russian language is banned in Ukraine?

**No, the statement about the ban of the Russian language in Ukraine is incorrect.**

\* Translated from Russian to English using Chrome build-in translation

# What's next?

 **Elon Musk**   [@elonmusk](#) Subscribe  

We are building Grokikipedia [@xAI](#).

Will be a massive improvement over Wikipedia.

Frankly, it is a necessary step towards the xAI goal of understanding the Universe.

 **David Sacks**   [@DavidSacks](#) · Sep 29

Wikipedia is hopelessly biased. An army of left-wing activists maintain the bias and fight reasonable corrections. Magnifying the problem, Wikipedia often appears first in Google search results, and now it's a trusted source for AI model training. This is a huge problem. [x.com/TuckerCarlson/](#)

1:49 PM · Sep 30, 2025 · **23.6M** Views

**611 million:** The number of people who logged into **X** at least once in a month. (2025)

 **Elon Musk**   [@elonmusk](#) · Oct 20  

Postponing **Grokikipedia** v0.1 launch to end of week.

We need to do more work to purge out the propaganda.

 11K  13K  140K  17M  

# Thank you!

*Presented by:*  
Mykola Trokhymovych  
[trokhymovych.com](http://trokhymovych.com)