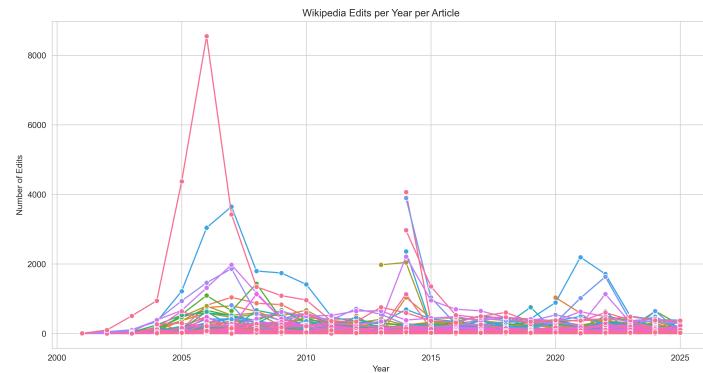


# Results – Midterm Analysis (28 Oct. 2025)

## 1. Motivation

In the file **Wikipedia\_articles.txt**, we gathered around **340 Wikipedia articles** related to **Ukrainian and Russian culture, history, and politics**.

By retrieving all the edits made to these articles, we aim to identify **patterns of activity** — for instance, peaks in editing activity during major historical or geopolitical events.



- 2014 Ukrainian presidential election
  - 2014 pro-Russia unrest in Ukraine
  - 2022 Russian invasion of Ukraine
  - Abortion in Ukraine
  - Ethnic groups in Ukraine
  - Empire of Ukraine
  - Alexander II of Russia
  - All-Ukrainian Union of Churches of Evangelical Christian Baptists
  - Andriy Shevchenko
  - Andriy Skurkovsky
  - Annexation of Crimea by the Russian Federation
  - Armed Forces of Ukraine
  - Armenian Diocese of Ukraine
  - Arts of Ukraine
  - Banking in Ukraine
  - Baptism of Kiev
  - Geography of Ukraine
  - Government of Ukraine
  - History of Christianity to the Orthodox Church of Ukraine
  - HIV/AIDS in Ukraine
  - Holodomor
  - Bessarabia
  - Bibliography of Ukrainian history
  - Books of the Eastern Baptist Churches and Ministries of Ukraine
  - Buddhism in Ukraine
  - Bukovina
  - Business in Ukraine
  - COVID-19 pandemic in Ukraine
  - Carpathian Ruthenes
  - Catholic Church in Ukraine
  - Censuses in Ukraine
  - Christianity in Russia
  - Christmas in Ukraine
  - Christmas Eve/Ukraine
  - Christmas in Ukraine
  - Crimean Tatars
  - Crimean Tatars in Ukraine
  - Communist Party of Ukraine (Soviet Union)
  - Communist Party of the Soviet Union
  - Corruption in Ukraine
  - Cossack Hetmanate
  - Crime in Ukraine
  - Crimean War
  - Crimean Karabakh
  - Crimean Khanate
  - Crimean Tatar language
  - Crimean Tatar literature
  - Crimean Tales
  - Crimean peninsula
  - Culture of Ukraine
  - Declaration of Independence of Ukraine
  - Directorate of Security of Ukraine
  - Dissolution of the Soviet Union
  - Dynasty of Ukraine
  - Dnipro Cossacks
  - Donets Basin
  - Eastern Front (World War I)
  - Eastern Front (World War II)
  - Economy of Ukraine
  - Education in Ukraine
  - Elections in Ukraine
  - Encyclopedia of Ukraine
  - Energy in Ukraine
  - English Civil War
  - Epiphanius I of Ukraine
  - Ethnic groups in Ukraine
  - Empire of Ukraine
  - Evangelical Lutheran Church in Russia, Ukraine, Kazakhstan and Central Asia
  - Evangelical Presbyterian Church of Ukraine
  - Evangelical Protestant Church of Ukraine
  - Evgenii Vaynshteyn
  - Flag of Ukraine
  - Folklore of Ukraine
  - Football in Ukraine
  - Foreign policy of Ukraine
  - Freedom of the press in Ukraine
  - Galicia (Central Europe)
  - Galician (Ukrainian)
  - Gender inequality in Ukraine
  - Geography of Ukraine
  - Government of Ukraine
  - History of Christianity to the Orthodox Church of Ukraine
  - HIV/AIDS in Ukraine
  - Health in Ukraine
  - Hills of Ukraine
  - History of the Greek Church
  - Hutsul language
  - Hinduism in Ukraine
  - Holocaust in present-day Ukraine
  - History of Christianity in Ukraine
  - History of Crimea
  - History of Dnipro Cossacks
  - History of the Catholic Church in Ukraine
  - History of Ukraine/Early modern period
  - History of Ukraine/World War II and the Nazi Occupation
  - History of Ukraine/Modern period
  - History of banking in Ukraine
  - History of jewellery in Ukraine
  - History of railroads in Ukraine
  - History of rugmaking in Ukraine
  - History of technology in Ukraine
  - History of television in Ukraine
  - History of the Jews in Ukraine
  - History of the Jews in Ukraine
  - History of the Soviet Orthodox Church
  - History of the Soviet Union (1953–1991)
  - History of the Soviet Union (1991–present)
  - Human rights in Ukraine
  - Human trafficking in Ukraine
  - Human rights in Ukraine
  - Hutsul language
  - Hutsul literature
  - Immigration to Ukraine
  - Imperial Russian Army
  - Imperial Russian Heritage of Ukraine
  - Invasion of Poland
  - Islam in Ukraine
  - Ivan Franko
  - Janosik
  - Judiciary of Ukraine
  - Kholmogory
  - Khmelnytsky Uprising
  - Khmelnytsky Governorate (Ukraine)
  - Kiev
  - Kievan Rus'
  - Kingdom of Galicia and Lodomiria
  - Kuban Cossacks
  - Kyiv
  - Kyiv Pechersk Lavra
  - LGBT rights in Ukraine
  - Languages of Ukraine
  - Latin Church in Ukraine
  - Law of Ukraine
  - Law of Ukraine
  - Law of Ukraine
  - Leonid Kravchuk
  - Leonid Kuchma
  - Lorraine cross (instrument)
  - List of Intangible Cultural Heritage elements in Russia
  - List of Intangible Cultural Heritage elements in Ukraine
  - List of Metropolitan and Patriarchs of Kiev/Metropolitan of Kiev and All Ukraine
  - List of Monuments of Archaeological and Historical Value of Ukraine
  - List of UNESCO World Heritage Sites in Ukraine
  - List of Ukrainian composers
  - List of Ukrainian folk songs
  - List of World Heritage Sites in Ukraine
  - List of castles in Ukraine
  - List of cities in Ukraine
  - List of cities in Ukraine
  - List of fortresses in Ukraine
  - List of longest rivers of Ukraine
  - List of mountains in Ukraine
  - List of museums in Ukraine
  - List of national parks of Ukraine
  - List of places named after people in Ukraine
  - List of waterfalls in Ukraine
  - List of world religions
  - Mass media in Ukraine
  - Mass media of Ukraine
  - Mass media in Poltava and Eastern Galicia
  - Media in Ukraine
  - Media of Ukraine
  - Ministry of Internal Affairs (Ukraine)
  - Ministry of National Defense (Ukraine)
  - Modern history of Ukraine
  - Motovilikh in Ukraine
  - Museums in Ukraine
  - Muslim culture in Ukraine
  - Mykhailo Hrushevsky
  - Mykola Kostomarov
  - Mykola Voronyi
  - Name of Ukraine
  - National Art Museum of Ukraine
  - National Bank of Ukraine
  - National Library of Folk Architecture and Folkways of Ukraine
  - National Parliamentary Library of Ukraine
  - National anthem of Ukraine
  - National emblem of Ukraine
  - Neo-Byzantine architecture in the Russian Empire
  - New Russia
  - Numbered Uprisings
  - Odesa
  - Opera in Ukraine
  - Orange Revolution
  - Organization of Ukrainian Nationalists
  - Orthodox Church of Ukraine
  - Outline of Ukraine
  - Outline of the Russo-Ukrainian War
  - Outline of World War I
  - Photography in Ukraine
  - Podolia
  - Poetry in Ukraine
  - Polish areas annexed by the Soviet Union
  - Polish-Soviet War
  - Political parties in Ukraine
  - Politics of Ukraine
  - Pop music in Ukraine
  - Popular places in Ukraine
  - Popular places in Ukrainian Rural settlements
  - Population transfers in the Soviet Union
  - Post-Soviet Union in Ukraine
  - President of Ukraine
  - Principality of Kiev
  - Princely court in Kiev
  - Protestantism in Ukraine
  - Protestants in Ukraine
  - Public music in Ukraine
  - Public places in Ukraine
  - Reichskommissariat Ukraine
  - Religion in Ukraine
  - Return of Dignity
  - Right-Bank Ukraine
  - Rock music in Ukraine
  - Roman Catholicism in Ukraine
  - Romanians in Ukraine
  - Russia
  - Russian Architecture
  - Russian Civil War
  - Russian Empire
  - Russian Memoria
  - Russian Orthodox
  - Russian Orthodox Church
  - Russian Orthodox Old-Rite Church
  - Russian Revolution of 1917
  - Russian Revolution of Ukraine
  - Russians in Ukraine
  - Russification of Ukraine
  - Russian-Ukrainian War
  - Russo-Ukrainian War/Full-scale Russian invasion of Ukraine (2022–present)
  - Russophiles
  - Ruthenians
  - Saint Sophia Cathedral in Kiev
  - Saint Sophia Cathedral, Kyiv
  - Science and technology in Ukraine
  - Seven Days of War in Ukraine
  - Seven Natural Wonders of Ukraine
  - Seven Wonders of Ukraine
  - Seven Wonders of Ukraine/Seven Natural Wonders
  - Seven Wonders of Ukraine/Six Natural Wonders
  - Soviet Top League
  - Soviet Ukraine
  - Soviet Union
  - Soviet power
  - Sport in Ukraine
  - Sub-Carpathian Reformed Church
  - Suicide in Ukraine
  - Symbols of Ukraine
  - Taras Shevchenko
  - Telecommunications in Ukraine
  - Timeline of Ukraine
  - The Church of Jesus Christ of Latter-day Saints in Ukraine
  - The Ukraine Week
  - Tourism in Ukraine
  - Transport in Ukraine
  - Turks in Ukraine
  - USSR
  - USSR national football team
  - Ukraine after the Russian Revolution
  - Ukraine after the Soviet Union
  - Ukraine at the Second World War
  - Ukraine national football team
  - Ukrainian Autocephalous Orthodox Church
  - Ukrainian Greek Catholic Church
  - Ukrainian Baroque
  - Ukrainian Bible Society
  - Ukrainian Greek Catholic Church
  - Ukrainian Institute of America
  - Ukrainian Insurgent Army
  - Ukrainian Lutheran Church
  - Ukrainian National Chernobyl Museum
  - Ukrainian National Committee
  - Ukrainian Orthodox Church of the Holy Trinity (Kyiv-Pechersk)
  - Ukrainian Orthodox Greek Catholic Church
  - Ukrainian Premier League
  - Ukrainian geopolitics
  - Ukrainian SSR
  - Ukrainian Soviet Socialist Republic
  - Ukrainian State
  - Ukrainian Year of Independence
  - Ukrainian animation
  - Ukrainian architecture
  - Ukrainian art
  - Ukrainian border guard
  - Ukrainian baroque
  - Ukrainian cuisine
  - Ukrainian culture
  - Ukrainian dance
  - Ukrainian diaspora
  - Ukrainian embroidery
  - Ukrainian folk tales
  - Ukrainian folk music
  - Ukrainian folklore
  - Ukrainian historical regions
  - Ukrainian hymns
  - Ukrainian language
  - Ukrainian literature
  - Ukrainian traditional clothing
  - Ukrainian people
  - Ukrainian underground
  - Ukrainian traditional building traditions
  - Ukrainian wine
  - Ukrainian wreath
  - Ukraine
  - Ukraine in Estonia
  - Unification council of the Orthodox churches of Ukraine
  - Vernadsky National Library of Ukraine
  - Velykyi Liubin
  - Victor Yushchenko
  - Volhynia
  - Volhynian genocide
  - Volodymyr the Great
  - Volyn
  - Western Ukrainian Cossacks
  - Wildlife of Ukraine
  - Wine of Ukraine
  - World Network of Biosphere Reserves in Europe and North America/Ukraine
  - World War II
  - Zapozhzhian Cossacks
  - Zapozhzhian Cossacks

## Observed Peaks of Activity

Distinct peaks can be identified around several key periods:

- **2005–2008** → Post–Orange Revolution period, marked by Russo-Ukrainian gas disputes and general geopolitical tension. This era also coincides with Wikipedia's "golden age," with roughly **50,000 active users** at its 2007 peak.
  - **2014** → The **Russian annexation of Crimea**.
  - **2020–2021** → The **COVID-19 pandemic**, which likely increased online activity overall.

- **2022–2025 → The large-scale Russian invasion of Ukraine**, generating global attention and narrative conflict.

These observations motivate a deeper investigation into **editor behavior and content modification patterns**.

---

## 2. Research Structure

This midterm analysis is divided into two complementary parts:

### 1. Weaponised vs. Non-Weaponised User Analysis

A global overview of editing patterns among a sample of **2,454 unique users**, distinguishing between edits classified as *weaponising* or *non-weaponising*.

### 2. Weaponising Users — Fine-Grained Analysis

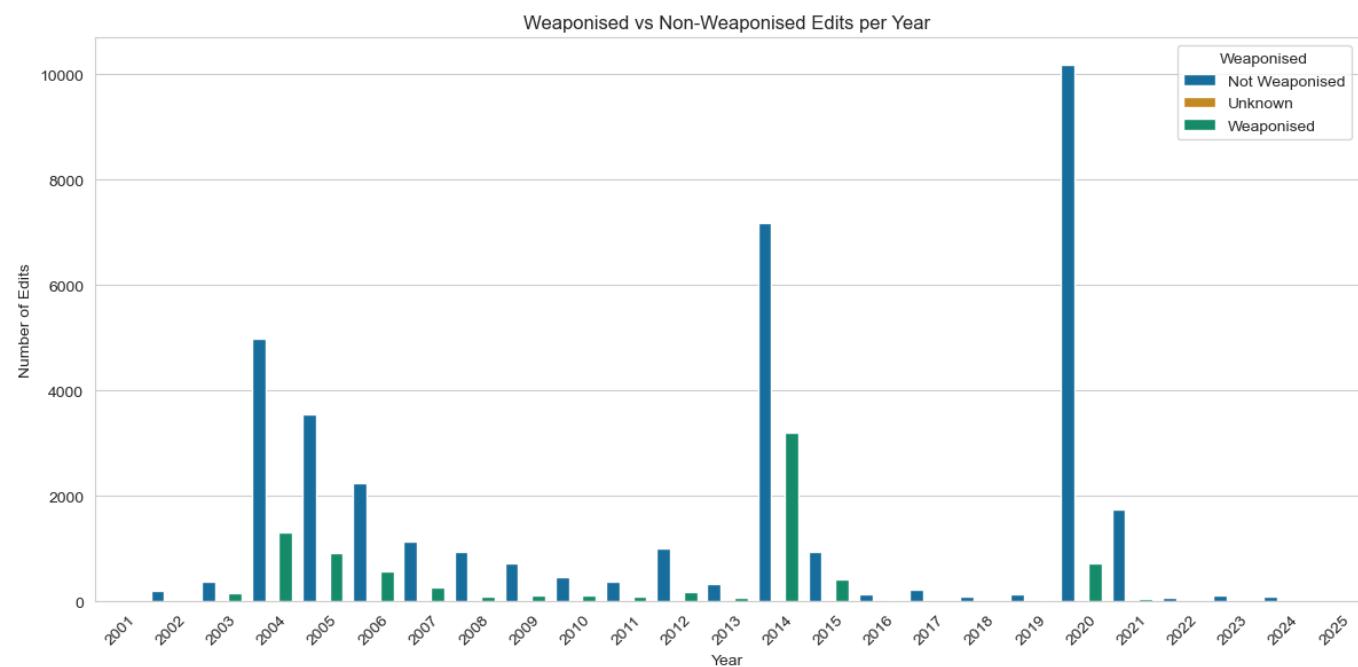
A focused investigation of **300 unique users** from the *fine-grained dataset*, restricted to edits identified as weaponising by the language model.

---

## 3. Weaponised vs. Non-Weaponised User Analysis

This section is based on the file **matched\_edits\_all.csv**, produced by **Emanuela**, which aligns fine-grained edits with corresponding user data from the *LLM Results* folder.

One can first check how *Weaponising VS Non weaponising* edits are distributed over time.



There is a clear correlation with the first graph, showing more edits on crucial geopolitical periods as well as a peak of weaponising edits. No peaks are displayed for the 2022–2025 periods since no article in the dataset was about Russian large scale invasion in Ukraine.

Metadata of the **matched\_edits\_all.csv**

```
{'total_edits': 45697,
'total_unique_users': 2454,
'number_unique_articles': 41,
'num_bot_users': 92,
'num_anonymous_users': 1011,
'num_registered_users': 1351,
'num_bot_edits': 1376,
'num_anonymous_edits': 8315,
'num_registered_edits': 36006,
'weaponised_edits': 8388,
'non_weaponised_edits': 37278,
'unknown_weaponised_edits': 31}
```

There is a lot we can analyse coming from this metadata. We first start to focus on the **registered** users and specially the top 10 most frequent editors among the 1351 registered users.

### Top 10 users general info

The top 10 registered editors are :

```
['LSGH',
'Toddy1',
'TheDraperyFalls',
'UP9',
'Genyo',
'Aleksandr Grigoryev',
'Dbachmann',
'Vasile',
'Irpen',
'Mzajac']
```

### Proportion regarding Ur/Ru articles

To identify how much each of the top 10 users focused on Ukraine- or Russia-related content, I filtered their edited articles using both small and large sets of keywords related to these two countries.

```
ukraine_keywords_small = [
    "ukraine", "ukrainian", "kyiv", "kiev", "crimea", "crimean", "kuban",
    "donbas", "donetsk", "luhansk",
    "maidan", "yanukovych", "yushchenko", "zelenskyy", "poroshenko",
    "catherine", "bukovina", "bessarabia", "eastern",
    "euromaidan", "dnipro", "odessa", "sevastopol", "putin", "rus",
    "russia", "russian", "moscow", "kremlin", "soviet"
]

ukraine_keywords_large = [
    # core country and people
    "ukraine", "ukrainian", "kyiv", "kiev", "crimea", "crimean", "kuban",
```

```
"donbas", "donetsk", "luhansk",
    "maidan", "yanukovych", "yushchenko", "zelenskyy", "poroshenko",
"catherine", "bukovina", "bessarabia",
    "eastern", "euromaidan", "dnipro", "odessa", "sevastopol", "putin",
"rus'", "russia", "russian", "moscow",
    "kremlin", "soviet",

    # politics & government
    "verkhovna rada", "president", "prime minister", "parliament",
"government", "cabinet", "federation",
    "referendum", "annexation", "independence", "revolution", "reforms",
"corruption", "sanctions",
    "occupation", "treaty", "agreement", "ceasefire", "negotiations",
"elections", "coup", "unification",

    # geography & regions
    "zaporizhzhia", "mariupol", "kharkiv", "kherson", "mykolaiv",
"chernihiv", "sumy", "poltava", "vinnysia",
    "lviv", "ivano-frankivsk", "ternopil", "lutsk", "uzhhorod",
"dniproptetrovsk", "donetsk oblast",
    "luhansk oblast", "transcarpathia", "prykarpattia", "galicia",
"novorossiya", "black sea", "azov sea",

    # historical references
    "kyivan rus", "tsar", "imperial", "empire", "ussr", "communist",
"lenin", "stalin", "bolshevik",
    "cold war", "perestroika", "glasnost", "collapse", "partition",
"catherine the great", "brezhnev",
    "chernobyl", "orange revolution", "revolution of dignity",
"holodomor", "soviet union",

    # war and military
    "invasion", "occupation", "annexed", "frontline", "offensive",
"defense", "army", "forces", "military",
    "russian troops", "ukrainian forces", "separatist", "rebels",
"paramilitary", "nato", "eu", "un", "war",
    "conflict", "shelling", "bombing", "airstrike", "occupation forces",
"mobilization", "martial law",

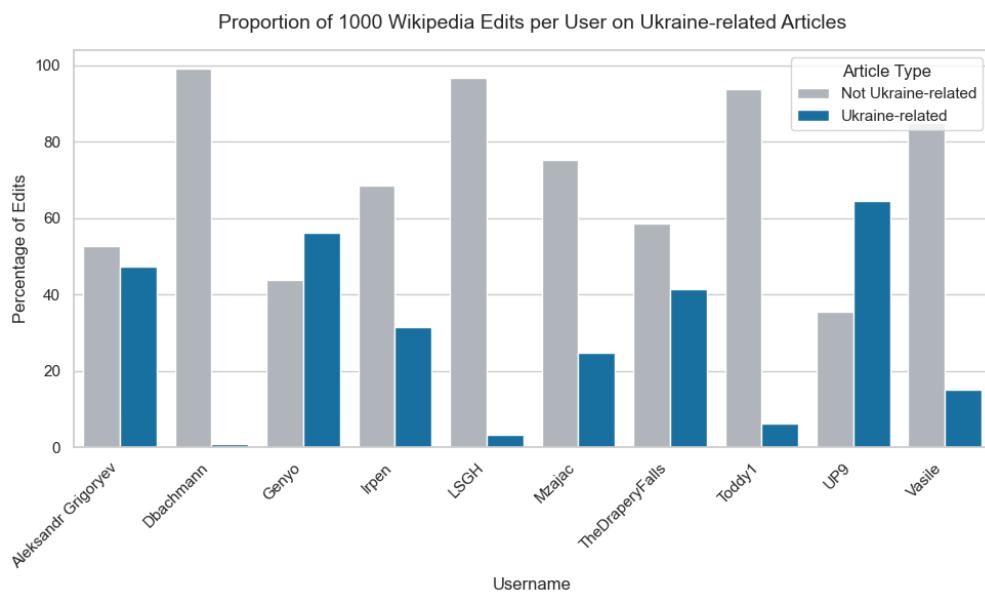
    # culture, identity & language
    "language", "identity", "heritage", "culture", "orthodox", "church",
"patriarch", "ukrainian language",
    "russian language", "minority", "bilingual", "autonomy",
"nationalism", "independence day", "flag",
    "anthem", "symbol", "national identity", "sovereignty",

    # current / modern references
    "donbas war", "russian invasion", "ukrainian front", "crimea bridge",
"moskva cruiser", "ukrainian army",
    "russian army", "zelensky", "kremlin propaganda", "occupation
administration", "territorial defense",
    "european union", "eu membership", "nato membership", "nato
expansion", "eu sanctions", "ukraine war",
    "full-scale invasion", "special military operation", "mobilisation",
```

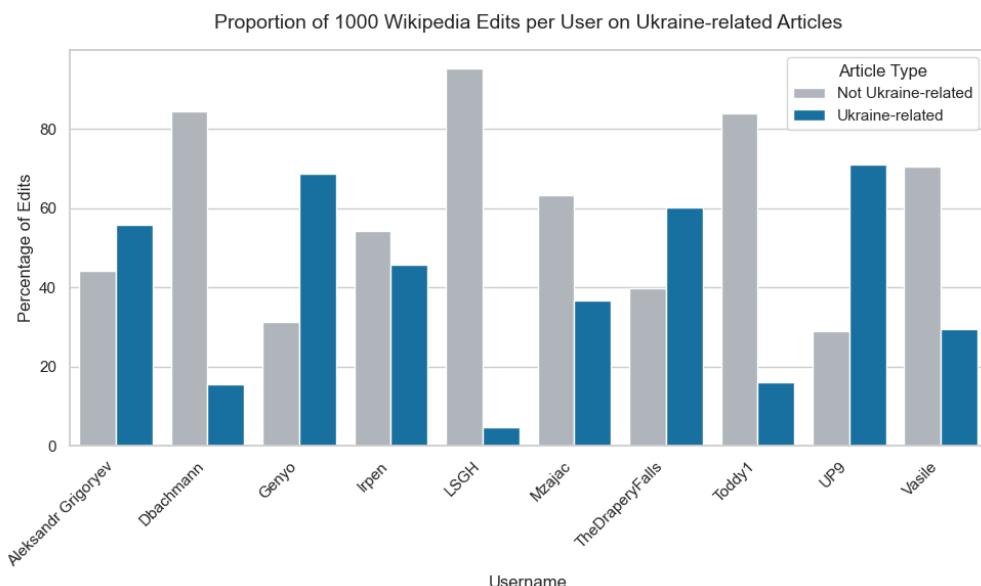
```
"referendum in crimea",
# other
"gas pipeline", "north stream", "energy crisis", "grain corridor",
"black sea fleet", "peace talks",
"donetsk people's republic", "luhansk people's republic", "kyiv
oblast", "liberation", "resistance",
"occupation zone", "ukrainian refugees", "mariupol steel plant",
"azovstal", "bucha", "irpin", "kharkiv offensive"
]
```

I did the analysis for a sub sample of 1000 analysis per user and for all the edits the user made since the creation of his account using Wikipedia's user API.

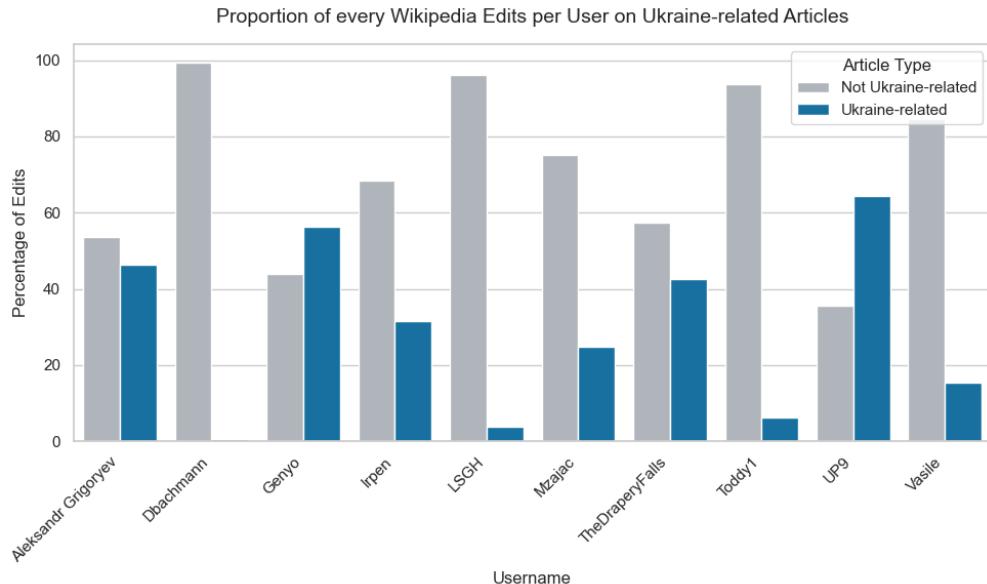
With the small set for 1000 edits per user :



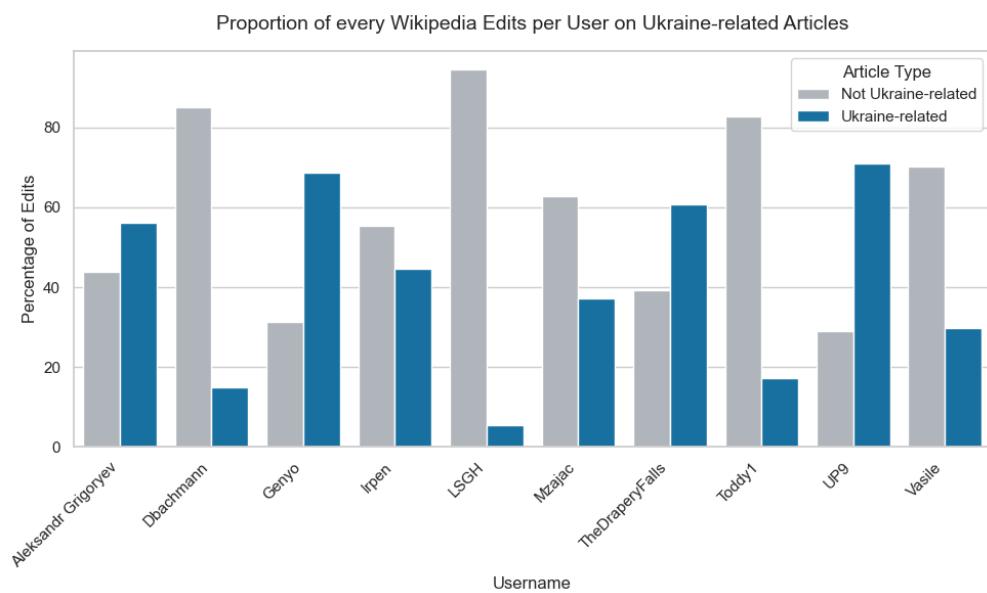
With the large set for 1000 edits per user :



With the small set for every edits per user :

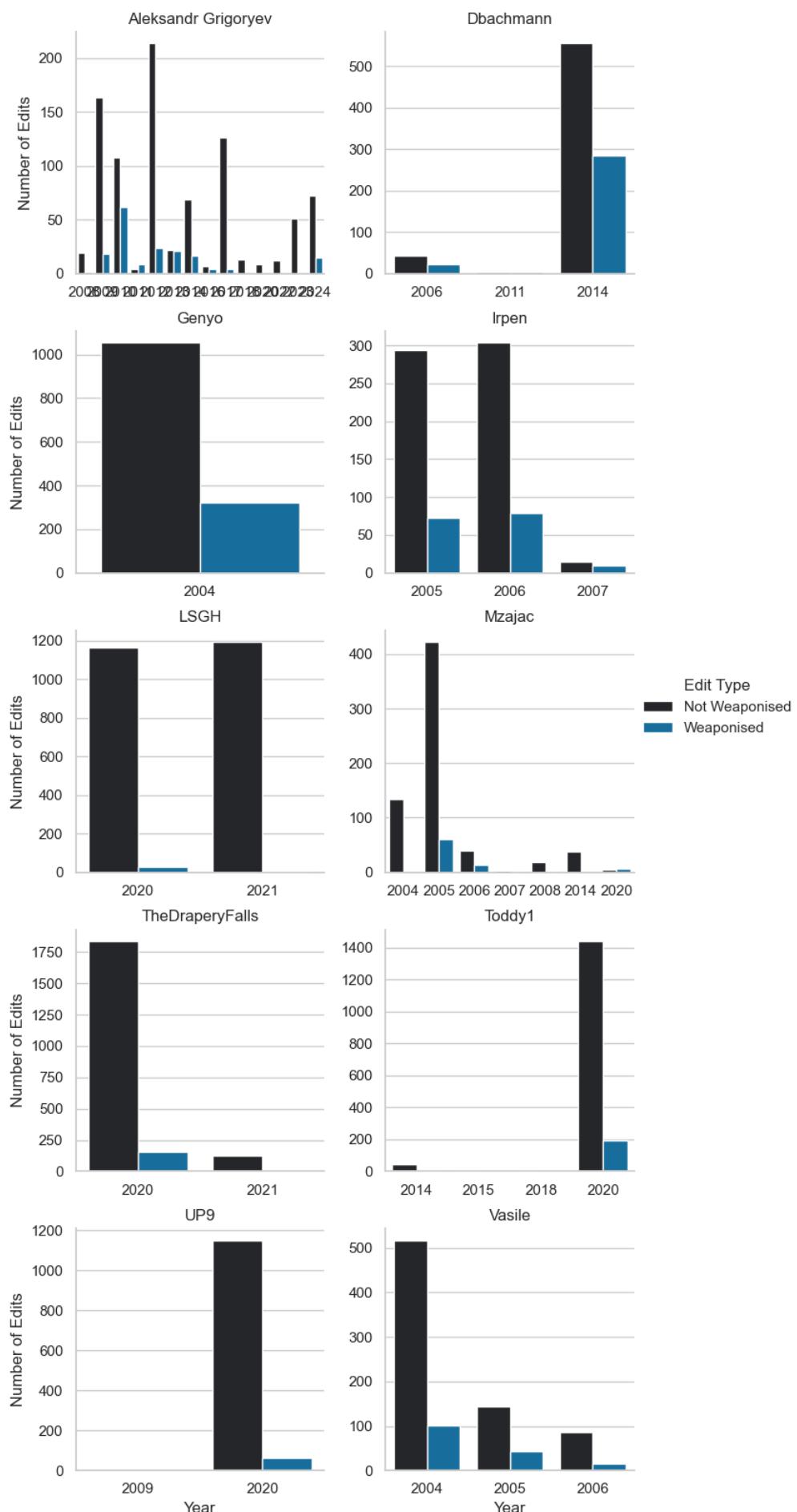


With the large set for every edits per user :



**Weaponising edits vs non weaponising edits per year for top 10 for only Ur/Ru related article**

### Weaponised vs Non-Weaponised Edits per Year — Top 10 Users



The graph shows that while Dbachmann seldom edits Ukraine- or Russia-related pages, those few edits are more likely to be weaponising compared to others. In general, these top users are broad Wikipedia contributors, with limited focus on Ukraine/Russia topics and little evidence of weaponising behaviour.

## Groups

Group	Who it includes	Typical criteria	Can edit semi-protected pages?	Can edit extended-protected pages?
*	Anyone (even IPs)	none	✗	✗
user	Registered users	account created	✓	✗
autoconfirmed	Older + $\geq 10$ edits	$\sim 4$ days, $\geq 10$ edits	✓	✗
extendedconfirmed	Sustained editors	$\geq 30$ days, $\geq 500$ edits	✓	✓

hard to get some relevant information for only the top 10 users. Maybe will be more relevant when focus only on the weaponising users for the fiengrained file.

## Registered vs IP address users

The file, as displayed in the metadata, includes 2,454 unique users, of which:

- 1,351 are registered accounts
- 1,011 are anonymous (IP addresses)
- 92 are bots

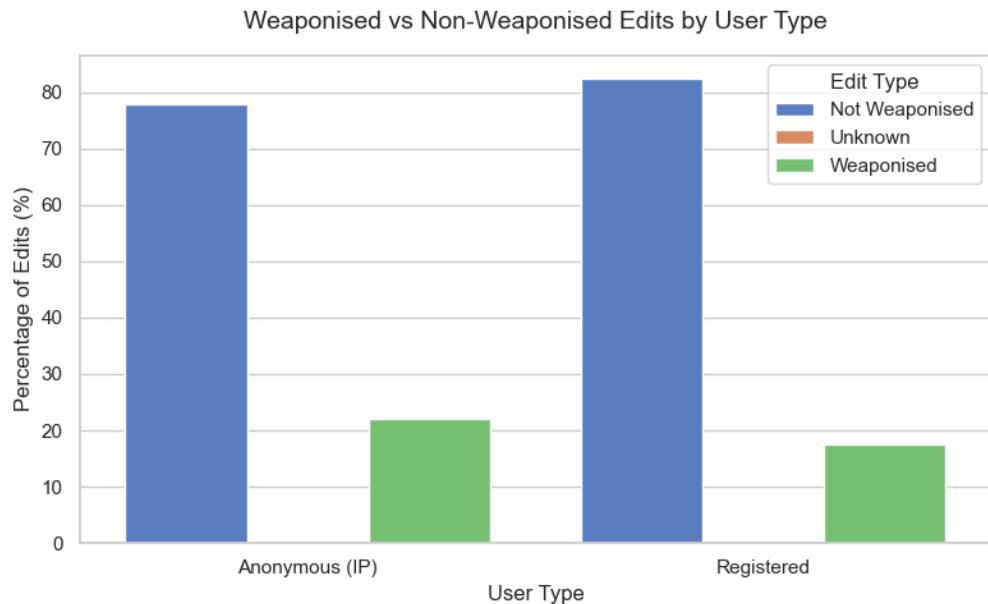
Although the number of registered and anonymous users is roughly comparable, the volume of edits differs significantly:

- Registered users: 36,006 edits
- Anonymous users: 8,315 edits

This implies that registered users contribute over four times more edits than anonymous users. Thus we can raise the following research question

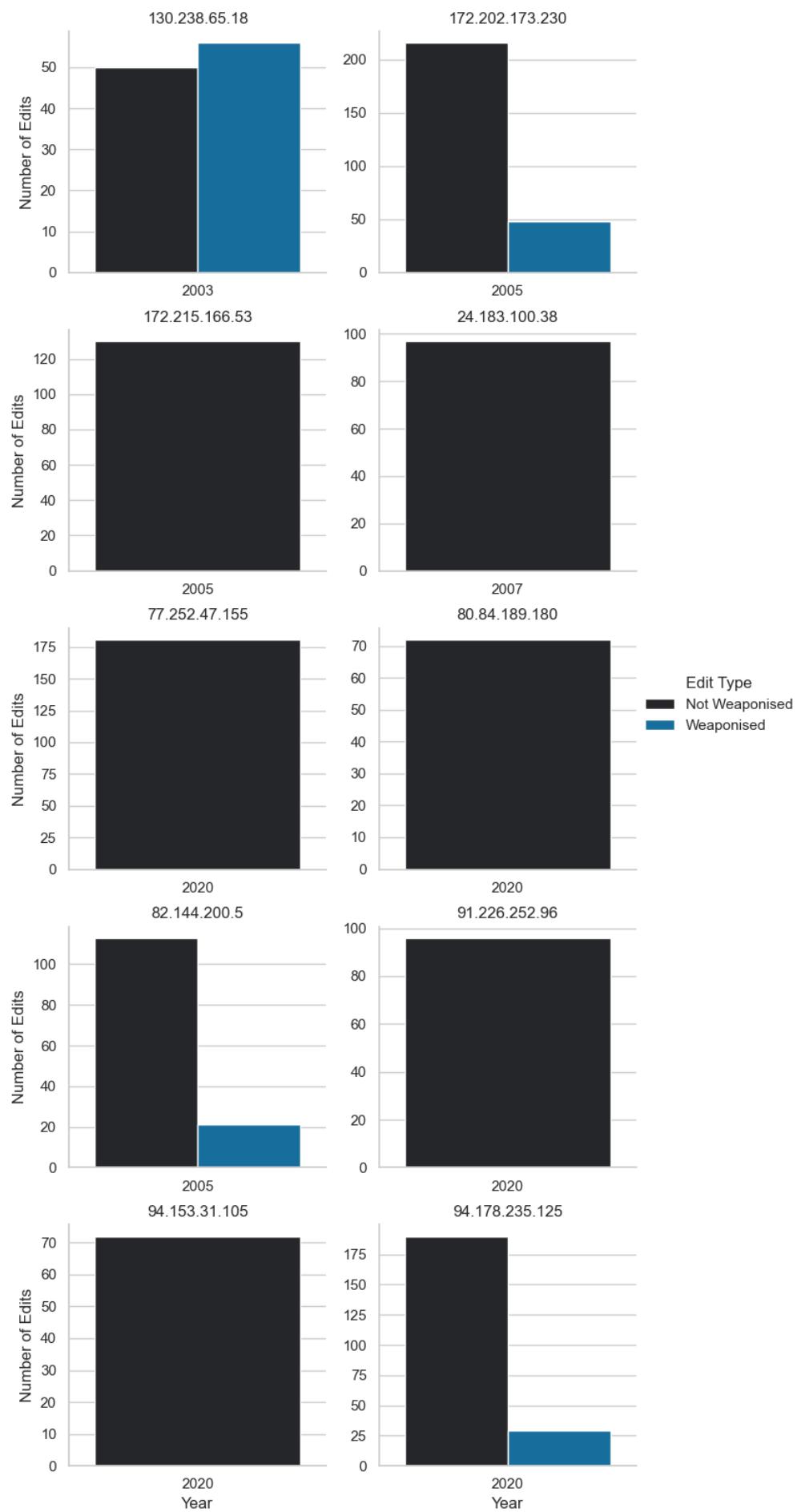
**Are anonymous (IP address) users more likely to produce weaponising edits compared to registered users due to their anonymity?**

The graph below shows that in proportion, IP address user are slightly more weaponising than registered user. They feel maybe more comfortable to weaponise under anonymity.



As we did for the top 10 registered users, we can also see how much edits and when the top 10 anonymous users are weaponising or not. As seen in the plot, we can see that the numbers of edits is way lower than the top 10 registered user and all of the top 10 IP adresses without exception are only editing during a single year, which is different than the registered users where the most of them are editing for a longer time frame.

## Weaponised vs Non-Weaponised Edits per Year — Top 10 Anonymous Users



By using an IP Geolocation API, we can check whether the used IP address is from a proxy or not and where the edit is located :

 [View Interactive IP Map](#)

---

### 3. Weaponised User Analysis

After analysing weaponised vs. non-weaponised edits for both registered and anonymous users, we can now go a step deeper and focus exclusively on the **users who have been classified as weaponising**. This allows us to explore not only *who* is weaponising, but also *how* different types of users (registered vs. anonymous) engage in specific forms of weaponisation.

**Note:** Unfortunately, I was not able to retrieve either the articles or the timestamps corresponding to the edits.

Metadata of the `matched_edits_all.csv`

```
{'total_edits': 513,
 'total_unique_users': 397,
 'num_bot_users': 9,
 'num_anonymous_users': 146,
 'num_registered_users': 242,
 'num_bot_edits': 11,
 'num_anonymous_edits': 156,
 'num_registered_edits': 346}
```

From the `fg_user_known.csv` file, we identify almost 400 weaponising users — including 146 anonymous (IP-based) editors and 242 registered editors.

Top 10 users general info

The top 10 registered editors are :

```
['Aleksandr Grigoryev',
 'RGloucester',
 'Yulia Romero',
 'Dbachmann',
 'Lvivske',
 'JASpencer',
 'Kuban kazak',
 'Irpen',
 'Mzajac',
 'Seryo93']
```

As expected, no IP addresses appear in this top 10. Some usernames have already appeared in the previous ranking, suggesting that a few users are consistently active across different subsets of edits. One can may prove those kind of behaviour in the next chapter "Graph".

To further characterise editing behaviour, we can ask:

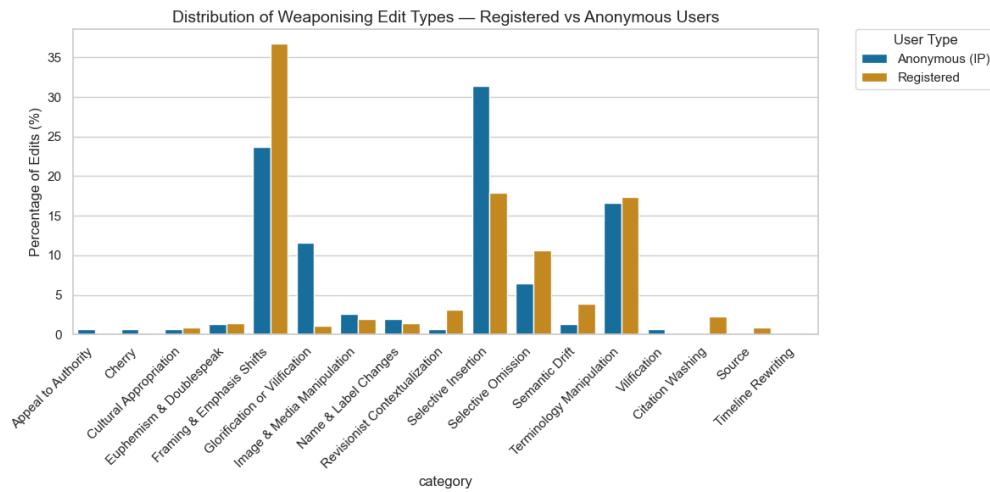
Do IP address editors are more concise and harmful than registered users ?

## Category of weaponisation per user type

This category is driven by the following RQ:

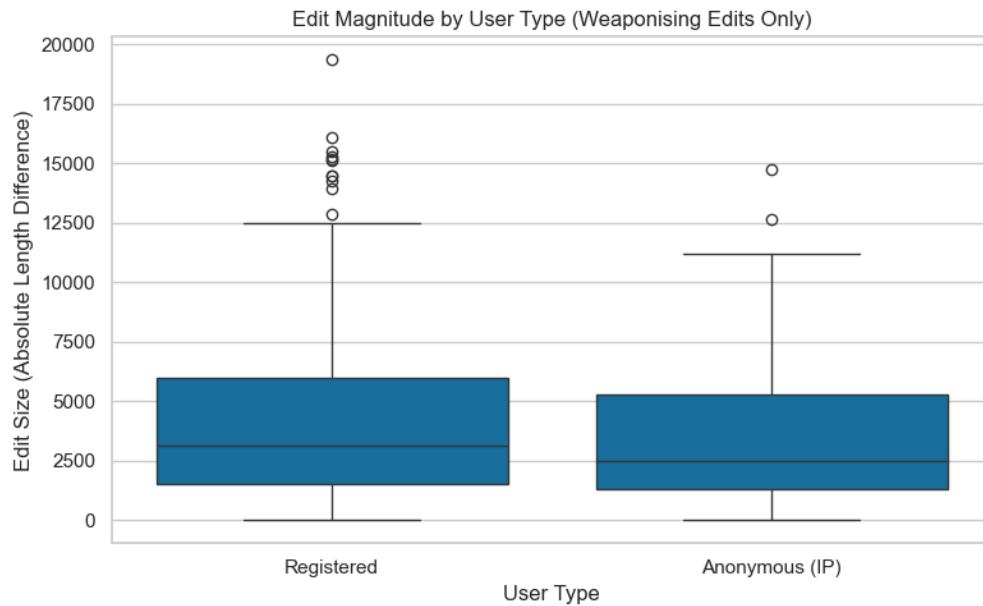
How do registered and anonymous users differ in the type of weaponising strategies they employ?

The graph below shows an interesting distribution of weaponisation categories across user types. On one hand, registered users appear to favour **Framing and Emphasis Shifts**\* : a technique that often involves adding context, rewording, or subtly reframing information to promote a particular point of view. These users generally produce longer edits and seem to embed manipulative intent in more sophisticated ways. On the other hand, anonymous users (IPs) tend to rely more on **Selective Insertion**, a strategy that typically consists of short, targeted, and often more direct interventions. This pattern suggests that anonymous users might engage in quicker, less nuanced, but potentially more disruptive forms of manipulation.



## A. Edit Magnitude

In order to emphasize this behaviour, one can perform a boxplot of edits magnitude per user type by checking the length of the `changed_version` vs the length of the `initial_version`:



The median weaponising edit size was comparable between registered and anonymous users, though registered users exhibited a greater variance and more extreme outliers. This suggests that while both groups engage in manipulative micro-edits (e.g., terminology shifts or short insertions), registered users occasionally perform larger structural interventions — possibly reframing or rewriting entire paragraphs. In contrast, anonymous users tend to contribute smaller, localized edits (consistent with quick, opportunistic interventions).

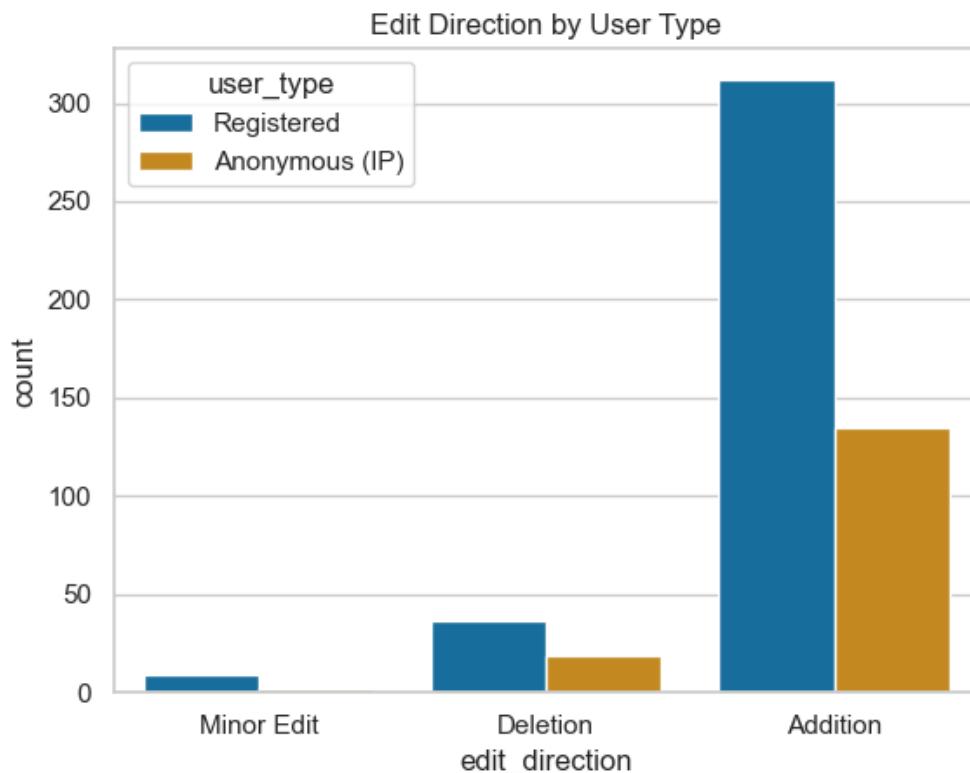
Maybe a good idea to assess the statistical significance of this results by calculating the p-value for both groups. So we suppose that the two groups does not provide a significal difference. By using Mann-Whitney U test i.e `mannwhitneyu` from `scipy.stats` module, we find a p-value of **0.03180**. This indicates that the observed difference between the two groups would occur by chance less than 3% of the time — providing moderate evidence of a statistically meaningful distinction in their editing behaviours.

## B. Change Direction

We can classify edits as Additions (positive  $\Delta\text{len}$ ), Deletions (negative  $\Delta\text{len}$ ), or Rewrites ( $\pm$  small  $\Delta\text{len}$ ) and compare distribution by user type. By folowing this question :

Are IPs more likely to add new sentences (insertions) or remove content (deletions)?

This graph shows no real



### C. Lexical & Semantic Analysis

Possible to analyse the sentiment polarity and toxicity by using model like [cardiffnlp/twitter-roberta-base-sentiment](#).

### D. Behavioral Significance

### ANNEX : Track the .csv files

```

all_user_profiles.csv           # Keeps the top10 user metadata gathered
with WP user API.
matched_edits_all.csv          # All the ngrams that matched with a
weaponising or non weaponising edits.
all_user_contribs_sub_n=1000.csv # Sub-sample of 1000 contributions of
top10 user gathered by WP user contrib API.
all_user_contribs.csv          # All contributions of top10 user
gathered by WP user contrib API.
df_user_known.csv               # All weap username matched in the
finegrained2
df_user_unknown.csv             # All the weap username not found in the
finegrained2
fg_user_known_delux.csv         # ll weap username matched in the
finegrained2 with extra info for analysis

```